

> 0 <
01 10 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-994-064-11.res made by shanley on Fri 27 Sep 102 18:57:03-PDT.

Query sequence being compared: US-09-994-064-11 (1-434)
Number of sequences searched: 1
Number of scores above cutoff: 1

Results of the initial comparison of US-09-994-064-11 (1-434) with:
File : seq67.pep

100-
N -
U 50-
M -
B -
E -
R -
O -
F 10-
S -
E 5-
Q -
U -
E -
N -
C -
E -
S 0-
SCORE 0 48 96 145 193 241 289 338 386 434
STDEV
PARAMETERS
Similarity matrix Unitary K-tuple
Mismatch penalty 6 Joining penalty 3
Gap penalty 5.00 Window size 30
Gap size penalty 0.33
Cutoff score 8
Randomization group 0

SEARCH STATISTICS

Scores: Mean 434 Median 0 Standard Deviation 0.00

Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 434
Number of sequences searched: 1
Number of scores above cutoff: 1

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name Description Init. Opt. Length Score Score Sig. Frame

1. US-09-994-064-67 Sequence 67, Application 434 434 434 0.00 0

1. US-09-994-064-11 (1-434)

US-09-994-064-67 Sequence 67, Application US/09994064

Initial Score = 434 Optimized Score = 434 Significance = 0.00
Residue Identity = 100% Matches = 434 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
MHRPHLRHSRYAKGEV LNKHMDCCGKRCCSGAAVFTLFWTCVRIMREHICFYRNAMDRHLFLRNAFWTIV
MHRPHLRHSRYAKGEV LNKHMDCCGKRCCSGAAVFTLFWTCVRIMREHICFYRNAMDRHLFLRNAFWTIV
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
LSSSFASOSTAAVTYDYILGRRALDAL TIPAVGPYRNLTRVSRGCDVVELNPI SNVDMMISAKEKEKGP
LSSSFASOSTAAVTYDYILGRRALDAL TIPAVGPYRNLTRVSRGCDVVELNPI SNVDMMISAKEKEKGP
80 90 100 110 120 130 140

150 160 170 180 190 200 210
FEASVWVFYIKGDGDEKRCPIYRKEYRECGDVQLSECAVQSAQMMAVDYVSTLVS RNAGLTFISPTA
FEASVWVFYIKGDGDEKRCPIYRKEYRECGDVQLSECAVQSAQMMAVDYVSTLVS RNAGLTFISPTA
150 160 170 180 190 200 210

220 230 240 250 260 270 280
ALSGQYLLTLKIGRFAQTALVTLLEVNDRLCKIGSQNLFLPSKCWTEQYQTFQGEHL YPIADTNRHADY
ALSGQYLLTLKIGRFAQTALVTLLEVNDRLCKIGSQNLFLPSKCWTEQYQTFQGEHL YPIADTNRHADY
220 230 240 250 260 270 280

290 300 310 320 330 340 350 360
YRGYEDILQRMNNLLRKKNPSPDPDPSPQEI PAVTKKAEGRT PDAESSEKKAPPEDESDMQAEASGEN
YRGYEDILQRMNNLLRKKNPSPDPDPSPQEI PAVTKKAEGRT PDAESSEKKAPPEDESDMQAEASGEN
YRGYEDILQRMNNLLRKKNPSPDPDPSPQEI PAVTKKAEGRT PDAESSEKKAPPEDESDMQAEASGEN
290 300 310 320 330 340 350 360

370 380 390 400 410 420 430
PALPEDEVPEDTEHDDPNSDPDYNDMPAVIPVEETTKSSNAVSMPIFAFVACAVALVGLVWSIVKCA
PALPEDEVPEDTEHDDPNSDPDYNDMPAVIPVEETTKSSNAVSMPIFAFVACAVALVGLVWSIVKCA
PALPEDEVPEDTEHDDPNSDPDYNDMPAVIPVEETTKSSNAVSMPIFAFVACAVALVGLVWSIVKCA
370 380 390 400 410 420 430

X
RS
RS
RS
X

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 22:36:32 ; Search time 2462.5 Seconds
(without alignments)
11090.002 Million cell updates/sec

Title: US-09-994-064-10
Perfect score: 1305
Sequence: 1 ATGCACCGTCTCATCTCAG.....TAAATGCGCGGTAGCTAA 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1305	100.0	1305	6	AR093367	AR093367 Sequence
2	1305	100.0	1305	6	AR135449	AR135449 Sequence
3	1305	100.0	18912	14	ILU28832	U28832 Infectious
4	1268.2	97.2	1627	14	HSMDGLYCO	L31965 Gallid herp
5	361	27.7	3502	6	A22122	A22122 ILTV gp60 g
6	52.2	4.0	161046	14	AF081810	AF081810 Lymantria
7	45.8	3.5	2676	1	MXARP0D	M32347 M.xanthus s
8	45.8	3.5	7980	1	MXU20669	U20669 Myxococcus
9	44.4	3.4	7218	6	I66494	I66494 Sequence 14
10	43.6	3.3	99360	8	NCB11N2	AL513444 Neurospor
11	43.4	3.3	125020	9	AF429315	AF429315 Homo sapi
12	43	3.3	3022	8	SCYOR053W	Z74961 S.cerevisia
13	43	3.3	54719	8	SCXV55KB	Z70678 S.cerevisia
14	41.6	3.2	390	6	AR135147	AR135147 Sequence
15	41.4	3.2	149333	2	AP004592	AP004592 Oryza sat
16	40.4	3.1	195549	2	AC012019	AC012019 Homo sapi
17	40	3.1	12204	1	AE005112	AE005112 Halobacte
18	40	3.1	23407	1	RSO245811	AJ245811 Ralstonia
19	40	3.1	30350	8	YSCN9315	U10398 Saccharomyc
20	40	3.1	197050	1	AL646081	AL646081 Ralstonia
21	40	3.1	341887	1	AP003006	AP003006 Mesorhizo
22	39.6	3.0	10293	1	AE005012	AE005012 Halobacte
23	39.6	3.0	73476	2	AC101502	AC101502 Mus muscu
24	39.6	3.0	162071	2	AC096932	AC096932 Rattus no
25	39.4	3.0	71517	9	AC073862	AC073862 Homo sapi
26	39.4	3.0	125020	9	AF429315	AF429315 Homo sapi
27	39	3.0	91452	2	AC104273	AC104273 Oryza sat
28	39	3.0	100974	2	AC104715	AC104715 Oryza sat
29	39	3.0	121544	2	AC092779	AC092779 Oryza sat
30	38.8	3.0	325	9	HSTNT5A	X79855 H.sapiens H
31	38.8	3.0	489	9	HSTNT5B	X79856 H.sapiens H
32	38.8	3.0	1141	6	AX083744	AX083744 Sequence
33	38.8	3.0	185327	9	AC007224	AC007224 Homo sapi
34	38.6	3.0	1405	8	SCCYC3G	X04776 Yeast CYC3
35	38.6	3.0	103682	8	SCU12980	U12980 Saccharomyc
36	38.6	3.0	163025	2	AC016366	AC016366 Homo sapi
37	38.6	3.0	181773	9	AC005881	AC005881 citb_79_e
38	38.6	3.0	189062	9	AL137025	AL137025 Human DNA
39	38.4	2.9	93354	9	AL133286	AL133286 Human DNA
40	38.4	2.9	161355	2	AC011154	AC011154 Homo sapi
41	38.4	2.9	164404	2	AC092075	AC092075 Oryza sat
42	38.4	2.9	173016	2	AC019123	AC019123 Homo sapi
43	38.2	2.9	476	3	TRBH1C8AAA	L27116 Trypanosoma
44	38.2	2.9	1314	3	AY051683	AY051683 Drosophila
45	38.2	2.9	30278	2	AC019782	AC019782 Drosophila

ALIGNMENTS

RESULT 1
AR093367 AR093367 1305 bp DNA linear PAT 08-SEP-2000
LOCUS Sequence 19 from patent US 6001369.
DEFINITION AR093367
ACCESSION AR093367
VERSION AR093367.1 GI:10020117
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1305)
AUTHORS Cochran,M.D. and Junker,D.E.
TITLE Recombinant fowlpox viruses and uses thereof
JOURNAL Patent: US 6001369-A 19 14-DEC-1999;
FEATURES Location/Qualifiers
source 1..1305
BASE COUNT 331 a 349 c 344 g 281 t
ORIGIN

Query Match	Score	DB	Length
100.0%;	1305;	6;	1305;
Best Local Similarity	100.0%;	Pred. No. 0;	

	Matches	1305;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	07
OY	1	ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACTACGCGAAAGAGAGGTGCTTAAC	60							
Dd	1	ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACTACGCGAAAGAGAGGTGCTTAAC	60							
OY	61	AAACACATGGATTGCGGTGGAACAACGGTGCTGCTCAGGCGCAGCTGTATTCACCTCTTTTC	120							
Dd	61	AAACACATGGATTGCGGTGGAACAACGGTGCTGCTCAGGCGCAGCTGTATTCACCTCTTTTC	120							
OY	121	TGCACTTGTCACGAGTATATCGCGGAGCATATCTGCTTTGTACGCAACGGTATGACC GC	180							
Dd	121	TGCACTTGTCACGAGTATATCGCGGAGCATATCTGCTTTGTACGCAACGGTATGACC GC	180							
OY	181	CATTATTTTTGAGGAATGCTTTTTGGACATACTGCTTTCTTCTTCCGCTAGCCAG	240							
Dd	181	CATTATTTTTGAGGAATGCTTTTTGGACATACTGCTTTCTTCTTCCGCTAGCCAG	240							
OY	241	AGCACCGCCCGCTCACGTACGACTACATTTTAGCGCCGTCGCCGCTCGACGGGCTAAC	300							
Dd	241	AGCACCGCCCGCTCACGTACGACTACATTTTAGCGCCGTCGCCGCTCGACGGGCTAAC	300							
OY	301	ATACCGCGGTTGGCCCCGTATAACAGATACCTCCTACAGGGTATCAAGAGGCTGGACGTT	360							
Dd	301	ATACCGCGGTTGGCCCCGTATAACAGATACCTCCTACAGGGTATCAAGAGGCTGGACGTT	360							
OY	361	GTCGAGCTCAACCCGATTTCTAAGCTGGACGACATGATATCGCGGCCCAAAGAAAAGAG	420							
Dd	361	GTCGAGCTCAACCCGATTTCTAAGCTGGACGACATGATATCGCGGCCCAAAGAAAAGAG	420							
OY	421	AAGGGGGCCCTTTCGAGGCTCCGTCGTCGTGTTCTACGTGATTAAGGGCGACGACGGC	480							
Dd	421	AAGGGGGCCCTTTCGAGGCTCCGTCGTCGTGTTCTACGTGATTAAGGGCGACGACGGC	480							
OY	481	GAGGACAAAGTACTGTCCAATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAACTG	540							
Dd	481	GAGGACAAAGTACTGTCCAATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAACTG	540							
OY	541	CTATCTGAATGCGCCGTTCAATCTGACACAGATGTGGCAGTGGACTATGTTCCTAGCACC	600							
Dd	541	CTATCTGAATGCGCCGTTCAATCTGACACAGATGTGGCAGTGGACTATGTTCCTAGCACC	600							
OY	601	CTTGATCGCGAAATGGCGGGGACTGACTATATTCCTCCCCACTGCTGCGCTCTGCG	660							
Dd	601	CTTGATCGCGAAATGGCGGGGACTGACTATATTCCTCCCCACTGCTGCGCTCTGCG	660							
OY	661	CAATACCTGCTGACCCCTGAAAAATCGGGAGATTTGCGCAAAACAGCTCTCGTAACTCTAGAA	720							
Dd	661	CAATACCTGCTGACCCCTGAAAAATCGGGAGATTTGCGCAAAACAGCTCTCGTAACTCTAGAA	720							
OY	721	GTTAAGATCGCTGTTTAAAGATCGGGTGCAGCTTAACCTTTTACCGTCAAAATGCTGG	780							
Dd	721	GTTAAGATCGCTGTTTAAAGATCGGGTGCAGCTTAACCTTTTACCGTCAAAATGCTGG	780							
OY	781	ACAACAGAAGATATCAGACTGGATTTCAAGCGGAACACCTTATCCGATCGCAGACACC	840							
Dd	781	ACAACAGAAGATATCAGACTGGATTTCAAGCGGAACACCTTATCCGATCGCAGACACC	840							
OY	841	AATACACGACACGGGAGCAGCTATATCGGGGATACGAAGATATTCTGCAGCGCTGGAAT	900							
Dd	841	AATACACGACACGGGAGCAGCTATATCGGGGATACGAAGATATTCTGCAGCGCTGGAAT	900							
OY	901	AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCCTCTCCAGATAGCGTCCGCAA	960							
Dd	901	AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCCTCTCCAGATAGCGTCCGCAA	960							
OY	961	GAAATTTCCCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGACGCAAGAAAGCAGCGAA	1020							
Dd	961	GAAATTTCCCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGACGCAAGAAAGCAGCGAA	1020							
OY	1021	AAGAAGGCCCTCCAGAGACTCGGAGGACGACATGCAGGAGAGGCTTCTGAGAAAAT	1080							
Dd	1021	AAGAAGGCCCTCCAGAGACTCGGAGGACGACATGCAGGAGAGGCTTCTGAGAAAAT	1080							

QY	1081	CCTGCCGCCCTCCCGAAGACGACGAGTCCCGCGAGGACACCGGACGACGATGATCCAAAC	1140
	1081	CCTGCCGCCCTCCCGAAGACGACGAGTCCCGCGAGGACACCGGACGACGATGATCCAAAC	1140
QY	1141	TCGGATCCTGACTATTACAATGACATGCGCGCGTGAATCCCGGTGAGGAGACTACTAAA	1200
	1141	TCGGATCCTGACTATTACAATGACATGCGCGCGTGAATCCCGGTGAGGAGACTACTAAA	1200
QY	1201	AGTTCTAATGCCGCTCCATGCGCCATATTCGCGCGCTGTAGCCTGCGCGGCTC	1260
	1201	AGTTCTAATGCCGCTCCATGCGCCATATTCGCGCGCTGTAGCCTGCGCGGCTC	1260
QY	1261	GTGGGCTACTGTTTGAGACATCGTAAATGCGCGCTGACTAA	1305
	1261	GTGGGCTACTGTTTGAGACATCGTAAATGCGCGCTGACTAA	1305
RESULT	2		
LOCUS	ARI35449	1305 bp	DNA linear PAT 16-JUN-2000
DEFINITION	Sequence 19 from patent US 6136318.		
ACCESSION	ARI35449		
VERSION	ARI35449.1	GI:14476121	
KEYWORDS	unknown.		
SOURCE	unknown.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1305)		
AUTHORS	Cochran, M.D. and Junker, D.E.		
TITLE	Recombinant fowlpox viruses and uses thereof		
JOURNAL	Patent: US 6136318-A 19 24-OCT-2000;		
FEATURES	Location/Qualifiers		
source	1..1305		
BASE COUNT	331 a 349 c 344 g 281 t		
ORIGIN			
Query Match	100.0%;	Score 1305;	DB 6; Length 1305;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1305;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGCACCGCTCCTCATCTCAGACGGCACTCGCGTTACTACGCGAAAGAGAGGCTTAAAC	60
	1	ATGCACCGCTCCTCATCTCAGACGGCACTCGCGTTACTACGCGAAAGAGAGGCTTAAAC	60
QY	61	AAACACATGGAATGCGGTGGAACACGGTGTGCTCAGCGCGACGCTGATTACCTCTTTC	120
	61	AAACACATGGAATGCGGTGGAACACGGTGTGCTCAGCGCGACGCTGATTACCTCTTTC	120
QY	121	TGACCTGTGTGTCAGGATTATGCGGAGACATATCTGCTTGTACGCAACGCTATGACCGC	180
	121	TGACCTGTGTGTCAGGATTATGCGGAGACATATCTGCTTGTACGCAACGCTATGACCGC	180
QY	181	CATTATTTTGTGAGGAATGCTTTTGGACTATCTGCTACTGCTTCTTCTTCGCTAGCCAG	240
	181	CATTATTTTGTGAGGAATGCTTTTGGACTATCTGCTACTGCTTCTTCTTCGCTAGCCAG	240
QY	241	AGCACCGCGCGCTCACGTCAGTACGACTACATTTTAGGCGGCTCGCGGCTCGACGCTAAC	300
	241	AGCACCGCGCGCTCACGTCAGTACGACTACATTTTAGGCGGCTCGCGGCTCGACGCTAAC	300
QY	301	ATACCGCGGCTTGCGCGTATTAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT	360
	301	ATACCGCGGCTTGCGCGTATTAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT	360
QY	361	GTCAGCTCAACCGGATTTCTAACGTGAGGACATGATATCGCGGCCCAAGAAAAAGAG	420
	361	GTCAGCTCAACCGGATTTCTAACGTGAGGACATGATATCGCGGCCCAAGAAAAAGAG	420
QY	421	AAGGGGCGCTTTCGAGGCTCCGTCGCTGCTTCTACGTGATTAGGGCGACGACGGC	480
	421	AAGGGGCGCTTTCGAGGCTCCGTCGCTGCTTCTACGTGATTAGGGCGACGACGGC	480

Db 421 AAGGGGGCCCTTTCGAGGCTCCGTCGTCGTGTTCTACGTGATTTAAGGGCGAGCGGC 480
OY 481 GAGGACAACTACTGTCCAACTATAGAAAAAGATGACAGGAATGTGGCAGCTACAACTG 540
Db 481 GAGGACAACTACTGTCCAACTATAGAAAAAGATGACAGGAATGTGGCAGCTACAACTG 540
OY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGGCACTGACTATGTTCCTAGCAC 600
Db 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGGCACTGACTATGTTCCTAGCAC 600
OY 601 CTTGTATCGCGAAATGGCGGGGAGCTGACTATATCTCCCCACTGCTGCGCTCTGCGC 660
Db 601 CTTGTATCGCGAAATGGCGGGGAGCTGACTATATCTCCCCACTGCTGCGCTCTGCGC 660
OY 661 CAATACCTGTGACCCCTGAAAAATCGGGAGATTGGCAAACAGCTCTGTAACCTTAGAA 720
Db 661 CAATACCTGTGACCCCTGAAAAATCGGGAGATTGGCAAACAGCTCTGTAACCTTAGAA 720
OY 721 GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACTTTTACCCTGAAATGCTGG 780
Db 721 GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACTTTTACCCTGAAATGCTGG 780
OY 781 ACAACAGAACAGTATCAGACTGGATTTCAGGGCGAACACCTTTATCCGATCGACAGACC 840
Db 781 ACAACAGAACAGTATCAGACTGGATTTCAGGGCGAACACCTTTATCCGATCGACAGACC 840
OY 841 AATACACGACACGCGGACGACGATATTCGGGGATACGAGATATTCGACGCGCTGGAAT 900
Db 841 AATACACGACACGCGGACGACGATATTCGGGGATACGAGATATTCGACGCGCTGGAAT 900
OY 901 AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCCCTGCTCCAGATAGCGTCCCGCAA 960
Db 901 AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCCCTGCTCCAGATAGCGTCCCGCAA 960
OY 961 GAAATTTCCCGCTGTAAACAGAAAGCGGAAGGGCGCACCCCGGAGCAGAAAGCAGCGAA 1020
Db 961 GAAATTTCCCGCTGTAAACAGAAAGCGGAAGGGCGCACCCCGGAGCAGAAAGCAGCGAA 1020
OY 1021 AAGAAGGCCCCCTCCAGAAAGCTCGGAGGACGACATGACAGCAGAGGCTTCTGAGAAAT 1080
Db 1021 AAGAAGGCCCCCTCCAGAAAGCTCGGAGGACGACATGACAGCAGAGGCTTCTGAGAAAT 1080
OY 1081 CCTGCCGCCCTCCCGAGACGACGAAGTCCCGGAGACACCGACGATGATCCAAAC 1140
Db 1081 CCTGCCGCCCTCCCGAGACGACGAAGTCCCGGAGACACCGACGATGATCCAAAC 1140
OY 1141 TCGGATCCTGACTATTACATGACATGCCCCCGCTGATCCCGGTGAGAGAGACTACTAAA 1200
Db 1141 TCGGATCCTGACTATTACATGACATGCCCCCGCTGATCCCGGTGAGAGAGACTACTAAA 1200
OY 1201 AGTTCTAATGCCGCTCCATGCCCATATTCGCGCGCTTCGTAGCCTGCGCGTCCGCTC 1260
Db 1201 AGTTCTAATGCCGCTCCATGCCCATATTCGCGCGCTTCGTAGCCTGCGCGTCCGCTC 1260
OY 1261 GTGGGGCTACTGTTTGAGCATCGTAAATGGCGCGTAGCTAA 1305
Db 1261 GTGGGGCTACTGTTTGAGCATCGTAAATGGCGCGTAGCTAA 1305

RESULT 3
LOCUS ILU28832 18912 bp DNA linear VRL 09-AUG-1996
DEFINITION Infectious laryngotracheitis virus US10, US2, protein kinase, UL47,
glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein
E, ORF9 genes, complete cds.
ACCESSION U28832
VERSION U28832.1 GI:1486484
KEYWORDS
SOURCE Infectious laryngotracheitis virus.
ORGANISM Gallid herpesvirus 1
REFERENCE 1 (bases 1 to 18912)
Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.

AUTHORS Wild,M.A., Cook,S. and Cochran,M.
TITLE A genomic map of infectious laryngotracheitis virus and the
sequence and organization of genes present in the unique short and
flanking regions
JOURNAL Virus Genes 12 (2), 107-116 (1996)
MEDLINE 97033380
REFERENCE 2 (bases 1 to 18912)
AUTHORS Wild,M.A.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1995) Martha A. Wild, Syntro Research Laboratory,
3535 General Atomics Court, San Diego, CA 92121, USA
FEATURES
source location/Qualifiers
1. 18912
/organism="Gallid herpesvirus 1"
/strain="USDA challenge strain"
/db_xref="taxon:10386"
<1. 2909
/note="large inverted repeat flanking the short unique
region"
/rpt_type=inverted
697. 1533
/codon_start=1
/product="US10"
/protein_id="AAC55093.1"
/db_xref="GI:1486485"
/translation="MEMLIDGCYPLALMDSDHITAHAVPGRGERRROGAVASESADS
VDPCIRIASRLMRELVIEISSELKDGYGEFTSARDRNALIAANERLSAFLGASRATR
GLGLRPWASTESVANSPTDPNNGLGLEAMEGIEGDFWLDSDLDREDESRTM
QSENMRFVIEKELISWLSRHLPADLASAERTSRLAAGHWCCLMHPRPCREACLYD
SIYVQSLFCVGTGRVPOSEMRREYLALRAGAAAAANSPEVASISIFARDAGIALALAR
RR"
complement(1900. 2784)
/note="coding region unique to avian herpesviruses"
/codon_start=1
/product="unknown"
/protein_id="AAC55094.1"
/db_xref="GI:1486486"
/translation="MSKCYCLARHLKSPRCVGRVAFEGGLATMSRPTSHLDLAFSA
AFRGTDLPGRFWRASQSCDIFFWPDLAAVYQAAARAYFEGERLGSJQVAEDITAND
PRIAPAAKRAVAAGVGLWTALSELVGGNGELESVKQIIPRAAWEIRDVPKVI
GPDILSFEFAVELPVLIRALRGAGHSRAHWNQSSAPAGLAIIRIGMEMVRSLLV
IALPLSNFTLPEDLPESQNSIRAFVAHLMNCVATDKIMSPDVRVPESEFSYSHCLRE
IIMCERAFICYPCNPPPKW"
2910. 16003
/note="short unique region"
complement(2916. 3605)
/codon_start=1
/product="US2"
/protein_id="AAC55095.1"
/db_xref="GI:1486487"
/translation="MAPVKVTIVSAVDSHYKLPNSRFELSDSGWKELVHAVKTMASYD
RPSTLSVIVRPASLYEVSSELFLPRMCRVIRFEGEGDPGVSPEPNSGLDAGFYHLS
SGAYAAKEFHVLVGLTADICMAALNLPAPKTFLLITETGKNFERGVEIFLVNGDKTTL
SLSHPSVWTTLAPSSLRTIPWYSTVKFLKVKPNSAAYCVSDSDGGERQPKFFLSLFK
SKKPRSPRRRR"
3694. 5124
/codon_start=1
/product="protein kinase"
/protein_id="AAC55096.1"
/db_xref="GI:1486488"
/translation="MRFRRICSRRAEKRRRTTENPLTSKRVCVLDSEFSRTMSLRPYA
EILPTAEGVERLAELVSVTMTERAEPVTENTAVNSIPRANENGONFAVAGDGPSTTEK
VDGSHTFDEASSDYAGVPPLAQTRLKHDEFLOHFRVLDDLVEGAVFICDVRYTE
EQORRGVNSTNOGSKCKRLAKYKNGTRAAOLENEILVLGRHNENVLKIQEIL
RYPDNTYMLTQRYQFDLYSYMYDEAFDWKDSMLKQTRIRIMQLSAVSIHKKLIH
RDIKLENIPLNCDGKTVLGDGFTVTPFENEREPEFEGWGVATVATNSPEILARDSYCEI
TDIWSGCVVLEMVNSHEFCPIGDGNGPHQQLKVIDSLVCDEDEFPDPNCNLNYLH
YASIDRAGHTVPSLIRNLHLPADVEYPLVKMLTFDWRLRPSAAEVLAMPFSAEEERT
ITIHGKHKPIRPEIRARVPRSMSEG"
5210. 7081
/codon_start=1
/product="UL47"
/protein_id="AAC55097.1"

/db_xref="GI:1486489"
 /translation="MTLPHRLTKRPFARFRCVSVFVTHYSETKLDRYKNKTMILYRPDST
 MRRSGGDANHRGIRPRRKSIGAFSAREKTGRNALTESSSSDMLDPFSTKEFGGKW
 TVDGPADITAEVLISQAMDVLQVHKHEDAEERVAESKPTIOPFNAPDGSWNAOD
 FTRAPIVYPSAEVLDAELKVGAFVSRVLOQVPTFRSKKSVTVRDAQSFLGDSFWRIM
 ONYTVVLRHOHITRLRHPSSKSIIVNCNDPLMYAYANQFHWGRMVRPSLKLASPEENI
 OHGPMAAVERNAGAGLELMPAMRAAEERDQTRLRLKALISLIDMDAAVLASPFYWRG
 VODTSREEPALGCLSEYFALVYLAEETVLATMDHALVEMRALADFNEDVDETRYID
 PVKNEYLNGAEGTLRGIVASNTALAVVCANTYSTIRKLPSVATSACNVAAYTETLKA
 RRGMSDIYRILQKEEFYIAMLQVATHANFCLNLIKRSVDTGAPFLFRASSEKRL
 QQLNMLCLPLVPIQYEDFSKAMGSELKREKLETFPKAISDRDPGSLRPLISDHAH
 ELIADGVREKPVIDEPRASVALSTAAGVKVKARRLTSVRAPVPGAGAVSARRKSEI"
 7245. .8123

```

/codon_start=1
/product="glycoprotein G"
/protein_id="AAC55098.1"
/db_xref="GI:1486430"
/translation="MSGFSNIGSIATVSLVCSLLICASYLGAPVLIDGLESPPEPGKII
IAQCNKRTIEVTVPMSDYSGRTEGVSEVKKWFYGNISNPESFVFGVDSEFTSGHEDLS
TCWALLIHNLNASVCRAADAGIPDEFDKQCEKVKQRLRSGVELGSYVSGNSLYLPGMY
DAGIYAYQLSVGKGKGYTSYLLDVGPNPGCHDQYGTYYSLADEASDLSYDVASPEL
DGMEEEDYSNCLDMPPLRPWTTCVCSHDVEEQENATDELYLWDEECAGPLDETYDERSEE
TTPRMVVEFSPSTLQ"
8333. .11290

```

/note="ORF5"
 /codon_start=1
 /protein_id="AAC55099.1"
 /db_xref="GI:1486491"
 /translation="MGTMVLRLRLLAVADAALPTGRFCRWVKVPPGCTIGENLAVLA
 ESPVTGHATYPPPEGAVSFOIPADPTLIRIRGATDEDLALERGTSADADNVFSL
 YRPRRIGHAYFTIGVATGOSTESSYVIRGLVNASLERSVRLTETPCDENFLONEP
 TWMSKRWLGPPSPYVRDNDVAVLTKAOYIGECYSNSAOTGLTSLNMTFFSPKRVN
 VTWTTGGPSPSRITVYSSRENGQVLRNVSDFLYKTTDPIDGRAMINYIANTSPADS
 GSVALFTAFREGKLPSAICDHRIDMSGTEPGEETTFEDCKMIETPYRAGSNVPRDD
 SIRPGATLPPFDTAAPDFDTGTSPTPTVPEPAITLLIPRSTDMGEFFSTARATGET
 LSVVQETDRTLSTPLTPTLPGESNTLFTPTAGISTETPSAAHETTCQSAETV
 VFTQSPSTESETARSQSQEPMWFTQTPSTEQALTOQIAETEAFLTQTPSAEQMTFT
 QTPGAETAPAQTPSTIPELFTQSRSTPETAARAPSAPEVFTQSSSTVEFTQTPS
 TYPKTTLSSTEPALFTRTQAGTEAFTQTSAAEPMTQSTHFTQAPSTVPKA
 TQTPSTESEVLQTQSPSTEPVPTRTIGAEPEITQTPSAPEVYTRSSSTMPETAQSTP
 LASQNPSTSSGTGTHNTEPRYPVQTPHTQKLYTENKTLSEPTVSEFHEMSTAESQTP
 PLIDVKIIVEKFSNDGEVATACVSTVKSPYRVEITWNVKVDLVDVMEISGNSAGVENS
 NEKQKQLYRYRTDGRTSVQIMCLICSTSHSPEDYCFDTSIIAREKDIAPLEYFTSDP
 QTAACITTLPSGVVPRFEWSLNNVSLPEYLTAITYVSHTAGQSTVWKSASARAGEAWIS
 GRGAGNIYECTVLISDGTRYVTRKRECLNTNTWIAVENGAQAQQLYSLSGLVSLGCSI
 SALVATLMTAIYF"
 11098.12402
 /codon_start=1
 /product="glycoprotein D"
 /protein_id="AAC55100.1"
 /db_xref="GI:1486492"

/translation="MHRPHLRHSRYIAKGEVILNKHMDCCGKRCCSGAAVFTLFWTCV
 RIMREHICFVRNAMDRHLFEUNAFWTIVLLSFASOSTAATVYDILGRADALATIP
 AVGYPNRYLTRVSRGCDYVELNPISNWDIMISAKKEKEGGPFEASVWVFYILKGDG
 EDKYCPYRKEYRECGDVLSECAVDSOMWAVYVSTLVSNGAGLTFFSPAL
 SGQYLLTLKIGRFAQTALVTLVENDRCLKIGSQLNPLPSKCWTEQYQTGEGEHLYP
 IADNTRHADDDVYRGYEDILQRWNNILRKNKNSPADBPDSVPQEIPAVTKAEGRT
 DAESESEKKAPPEDEEDMQAEASGENPALPEDEVEDTEHDDPNSDPDYYNDMPAV
 IPVEETTKSSNAVSMPLEFAAFVACAVALVGLVMSLVKCARs"
 12510. .13598

```

/codon_start=1
/product="glycoprotein 1"
/protein_id="AAC5101.1"
/db_xref="GI:1486493"
/translation="MASLIGTLALLAATLAPFGANGCIVITGNHVSARIDDDHIVAP
RPEATIQQLFEMPGORPHKPYSGTVAFASDLTNOQCYQLSEERFENCRRSSVF
VGCKVTEYTFSSASNRLTGPPHPEFLTIRNRPNDSGMEYIVRLDDTKETIDVFIQL
SVYQFANTATATRGLYSKASCRTGFKLTLPVLEAYLTETESWRNMQAVATEPTTSAEA
TTPTPVATASASELAEHFTFPWLENGVDHYEPTPANENSNVTVRLGTMSPLLIGTV
AAVVSATIGLIVIVISIVTRNMCQPHRLKIDTVSQDDEERSQTRRESRKFGPMACEINK
GADQDSELVELVAIVNPSALSSPDSIKM"
13792. .15291
/codon_start=1

```

```

/product="glycoprotein E"
/protein_id="AAC5102.1"
/db_xref="GI:1486494"
/translation="MNMIVIVLASCLARLTFTATHVLELEGTQAVLGEDDPRNVEGT
VIKWTKVLRNACCKMAADVCSSPNYCFHDLIYDGKKDCPPAGPLSANLVIILKREGS
EYVLGSGGLHNSNITINIMTETREGLLFDPVTRSDGEIYFRIISQPLAMETTSYNVSL
SHVDEKAPAPHEVEIDITIKPSEAAHVELQMLPEHELDNDNSPTIYPLRYVEPPTAA
KFNVTYSWYGFEDVKEECEEEKLFEPVCYHPTDRCQFPATNQRLIGSLVMAEFLGAA
SLDCCSRDITLEDCHENRYNPLRFSRLSESAGALISPLAIPKVLIIVSDGILGW
STYVLGKRNSPRVVEVETHMPSKVPMNKVIGSPGMDETGNKMYFVAVGAATCVIL
TCALLVGKKKCPAHQMGTFESTKEPLIYAPLPKNEEAGGLTDDEVIYDEYVEPLFRGY
CKQEFREDVNTFFGAVVEGERALNFKSAIASMDRLIANKSGRRNMDSY"
15298..16080
/cds
/feature="ORF9"
/codon_start=1
/protein_id="AAC5103.1"
/db_xref="GI:1486495"
/translation="MPFKTGAEDAAAGKNREKKSRRNEILPTRLRGTKKTAGLSNY
TQPIPMWPKFCSARGESDNHACKDTFYRRTCCASRSTVSSQPDSPHTPMDTEYGRVPS
ARRKKLSSSDCEGAHQPLVSCKLPLDSQAAPARTSSAQRITVDEVSSPPPGVDAVA
LETRAELPGATTEQTESKNKLPLNQOSRLKPKPTNHHVGGERCPSGEIVAPSLIGILS
VGAIANELARMRACIPLAASAAAGIVAWAARALOKOGR"
16004..>18912
repeat_region

```

```

/note="large inverted repeat flanking the short unique
region"
/rpt_type=inverted
16129..17013
/note="coding region unique to avian herpesviruses"
/codon_start=1
/product="unknown"
/protein_id="AAC55104.1"
/db_xref="GI:1486496"
/translation="MSKCYCLARHLKYSPRCVGRVVAFGGLATMSRPTSHDLAESA
AFRGTDLPGGRFWRASQDIFFWPDLAIVQARAYFEKGERLGLQVAEDITAH
PRAPAKKRAVAAVAAGSLMIALSELVGGPNGELESKVWGKIIPRAAAMEIDVPKVPVI
GPILSFESAVALPVLVIRARGAHSRSAHMNNQSSAPAGLAIIRIGEMVRSLLV
IALPLSNFTLPEDLPEGSONSIRAFVAHLMNCVATDKIMSPDVRVPVEESFYSHCLRE
IIMCERAFQYPCNPPPKW"
complement(17380..18216)
/codon_start=1
/product="US10"
/protein_id="AAC55105.1"

```

Query Match	Best Local Similarity	Score	DB	Length
100.0%;	100.0%;	1305;	14;	18912;
Pred. No. 0;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 1305;				
QY 1	ATGCACCGTCTCATCTCAGACGGCAGCTCGCTTACTACGCCGAAAGGAGAGGTGCTTAAC	60		
DB 11098	ATGCACCGTCTCATCTCAGACGGCAGCTCGCTTACTACGCCGAAAGGAGAGGTGCTTAAC	11157		
QY 61	AAACACATGATTTGCGGTGGAAGCGTCTGCTCAGGGCAGCTGTATTCACTCTTTTC	120		
DB 11158	AAACACATGATTTGCGGTGGAAGCGTCTGCTCAGGGCAGCTGTATTCACTCTTTTC	11217		
QY 121	TGGACTTGTGTACAGATTATGCGGGAGCATATCTGCTTTGTACGCAAGGCTATGAGCCG	180		
DB 11218	TGGACTTGTGTACAGATTATGCGGGAGCATATCTGCTTTGTACGCAAGGCTATGAGCCG	11277		
QY 181	CATTATTTTGTAGSAATGCTTTTGGACTATCGTACTGCTTCTTCTTCGCTAGCCAG	240		
DB 11278	CATTATTTTGTAGSAATGCTTTTGGACTATCGTACTGCTTCTTCTTCGCTAGCCAG	11337		
QY 241	AGCACCGCGCCGCTACAGCTACGACTACATTTTAGGCCGTCGGCGCTCGACGCGCTAAC	300		
DB 11338	AGCACCGCGCCGCTACAGCTACGACTACATTTTAGGCCGTCGGCGCTCGACGCGCTAAC	11397		
QY 301	ATACCGGCGGTTGGCCGTATTAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT	360		
DB 11398	ATACCGGCGGTTGGCCGTATTAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT	11457		
QY 361	GTCAGCTCAACCCGATTTCTAACGTGAGAGACATGATATCGGGCGCCAAAGAAAAAGAG	420		

```

Db 11458 GTCGAGCTCAACCCGATTCTTAACGTGGACGACATGATATCGCGGCCCAAGAAAAAGAG 11517
QY 421 AAGGGGGCCCTTTGAGGCGCTCCGCTGCTGTTCTACGTGATTAAGGCGAGCAGCGC 480
Db 11518 AAGGGGGCCCTTTGAGGCGCTCCGCTGCTGTTCTACGTGATTAAGGCGAGCAGCGC 11577
QY 481 GAGGACAGTACTGTCCAATCTATAGAAAAAGAGTACAGGGAATGTGGCGAGCTACAAC 540
Db 11578 GAGGACAGTACTGTCCAATCTATAGAAAAAGAGTACAGGGAATGTGGCGAGCTACAAC 11637
QY 541 CTATCTGAATGCGCCGCTCAATCTGCACAGATGTGGCGAGTGCATATGTTCTTAGCACC 600
Db 11638 CTATCTGAATGCGCCGCTCAATCTGCACAGATGTGGCGAGTGCATATGTTCTTAGCACC 11697
QY 601 CTTGTATCGCGAAATGGCGCGGAGTACTATATCTCCCCACTGCTGCGCTCTGCGC 660
Db 11698 CTTGTATCGCGAAATGGCGCGGAGTACTATATCTCCCCACTGCTGCGCTCTGCGC 11757
QY 661 CAATACTTGTGACCCCTGAAAAATCGGGAGATTGCGCAACAGCTCTCGTAACCTAGAA 720
Db 11758 CAATACTTGTGACCCCTGAAAAATCGGGAGATTGCGCAACAGCTCTCGTAACCTAGAA 11817
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTCGCAGCTTAACCTTTTACCCTCGAATGCTGG 780
Db 11818 GTTAACGATCGCTGTTAAAGATCGGGTCGCAGCTTAACCTTTTACCCTCGAATGCTGG 11877
QY 781 ACAACAGACAGATATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGCAGACACC 840
Db 11878 ACAACAGACAGATATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGCAGACACC 11937
QY 841 AATACACGACACGCGGACGACGTATATCGGGGATACGAGATATTTCTGCAGCGCTGGAAT 900
Db 11938 AATACACGACACGCGGACGACGTATATCGGGGATACGAGATATTTCTGCAGCGCTGGAAT 11997
QY 901 AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCGCAA 960
Db 11998 AATTGCTGAGGAAAAAGATCCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCGCAA 12057
QY 961 GAAATTTCCCGCTGTAACCAAGAAACGGAAGGCGCACCCCGGAGCGCAGAAAAAGCGAA 1020
Db 12058 GAAATTTCCCGCTGTAACCAAGAAACGGAAGGCGCACCCCGGAGCGCAGAAAAAGCGAA 12117
QY 1021 AAGAAGGCCCTCCAGAAAGACTCGGAGGACGACATGACAGCGCAGAGGCTTCTGAGAAAAAT 1080
Db 12118 AAGAAGGCCCTCCAGAAAGACTCGGAGGACGACATGACAGCGCAGAGGCTTCTGAGAAAAAT 12177
QY 1081 CCTGCCGCCCTCCCGAGAGACGACGAAGTCCCGCGAGACACCCGAGCAGATGATCCAAC 1140
Db 12178 CCTGCCGCCCTCCCGAGAGACGACGAAGTCCCGCGAGACACCCGAGCAGATGATCCAAC 12237
QY 1141 TCGGATCCTGACTATTAACAATGACATGCCCCCGCTGATCCCGGTGAGGAGACTACTAA 1200
Db 12238 TCGGATCCTGACTATTAACAATGACATGCCCCCGCTGATCCCGGTGAGGAGACTACTAA 12297
QY 1201 AGTTCTAATGCCGCTCCATGCCCATATTCGGCGGCTTCGTAGCCTGCGCGGTGCGGCTC 1260
Db 12298 AGTTCTAATGCCGCTCCATGCCCATATTCGGCGGCTTCGTAGCCTGCGCGGTGCGGCTC 12357
QY 1261 GTGGGGCTACTGTTTGAGAGCATCGTAAAAATGCGCGGTAGCTAA 1305
Db 12358 GTGGGGCTACTGTTTGAGAGCATCGTAAAAATGCGCGGTAGCTAA 12402

```

```

RESULT 4
HSMGDGLXCO 1627 bp DNA linear VRL 06-MAY-1996
LOCUS HSMGDGLXCO 1627 bp DNA linear VRL 06-MAY-1996
DEFINITION Gallid herpesvirus 1 glycoprotein D (gd) gene, complete cds.
ACCESSION L31965
VERSION L31965.1 GI:493595
KEYWORDS glycoprotein D.
SOURCE Gallid herpesvirus 1.
ORGANISM Gallid herpesvirus 1.
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

```

```

REFERENCE 1 (bases 1 to 1627)
AUTHORS Johnson,M.A., Tyack,S.G., Prideaux,C.T., Kongsuwan,K. and Sheppard,M.
TITLE Sequence characteristics of a gene in infectious laryngotracheitis virus homologous to glycoprotein D of herpes simplex virus
JOURNAL DNA Seq. 5 (3), 191-194 (1995)
MEDLINE 95337426
FEATURES
    source
        1..1627
            /organism="Gallid herpesvirus 1"
            /strain="SA-2"
            /db_xref="taxon:10386"
            /note="also known as infectious laryngotracheitis virus (ILT)"
            144..151
                /gene="gd"
            144..1517
                /gene="gd"
            326..1441
                /gene="gd"
            /codon_start=1
            /product="glycoprotein D"
            /protein_id="AA98925.1"
            /db_xref="GI:493596"
            /translation="MDRHLFLRNAFWIYLLSSFASQSTAATVYDYLGRALDALTI
            PAVGPNRYLTRYRSKCELVLENPISVNDMISAKEKEKGGPEASVSEYVKGDD
            GEDKYCPYRKREYREGCDVOLLSECAVOSQMWALDYVSSSTLVSRNAGLTIFSPNA
            LSGQYLLTLKIGRFAOTLALVTLVLENDRLKIGRLNPLPSKWTTEQYOTGFOGELY
            PIADTNRHADVDVRYRVEDILQRWNNLLRKNPSAPDPRPDSVPOEIPAVTKKAEGRT
            PDAESSEKRAPDESSEDMDQAEASGENPALPEDEVPEDTEHDDPNSDPDYNDMPA
            VPIVEETKSSNAVSMPIFAAFVACAVASWGYWFGAS"
            1507..1517
                polYA_signal
                /gene="gd"
BASE COUNT 410 a 428 c 428 g 361 t
ORIGIN
Query Match 97.28; Score 1268.2; DB 14; Length 1627;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1287; Conservative 0; Mismatches 13; Indels 1; Gaps 1;
QY 1 ATGCACCGCTCCTCATCTCTACAGCGGCACCTCGCGTTACTACGCGGAAAGAGAGAGTGCCTTAAC 60
Db 155 ATGCACCGCTCCTCATCTCTACAGCGGCACCTCGCGTTACTACGCGGAAAGAGAGAGTGCCTTAAC 214
QY 61 AAACACATGATTGCGGTGGAACACGGTGTGCTCAGCGCGCAGCTGTATTCACCTTTTC 120
Db 215 GAACACATGATTGCGGTGGAACACGGTGTGCTCAGCGCGCAGCTGTATTCACCTTTTC 274
QY 121 TGGACTTGTGTAGGATATGCGGAGCATATCTGCTTGTAGCGAACGCTATGACCGC 180
Db 275 TGGACTTGTGTAGGATATGCGGAGCATATCTGCTTGTAGCGAACGCTATGACCGC 334
QY 181 CATTTATTTTGAGGAATGCTTTTGGACTATCTGCTTCTTCTTCCCTAGCCAG 240
Db 335 CATTTATTTTGAGGAATGCTTTTGGACTATCTGCTTCTTCTTCCCTAGCCAG 394
QY 335 CATTTATTTTGAGGAATGCTTTTGGACTATCTGCTTCTTCTTCCCTAGCCAG 394
QY 241 AGCACCGCGCGCTACAGTACATTTTAGCGCGCTCGCGGCTCGACCGCTTAAC 300
Db 395 AGCACCGCGCGCTACAGTACATTTTAGCGCGCTCGCGGCTCGACCGCTTAAC 454
QY 301 ATACCGGCGGTTGGCCGCTATAACAGATACCTCACTAGGGTATCAAGAGGCTGCGACGTT 360
Db 455 ATACCGGCGGTTGGCCGCTATAACAGATACCTCACTAGGGTATCAAGAGGCTGCGACGTT 514
QY 361 GTCGAGCTCAACCCGATTTCTAACGTGGAGCAGATGATATCGCGGCCAAGAAAAAGAG 420
Db 515 GTCGAGCTCAACCCGATTTCTAACGTGGAGCAGATGATATCGCGGCCAAGAAAAAGAG 574
QY 421 AAGGGGGCCCTTTGAGGCGCTCCGCTGCTGCTTCTACGTGATTAAGGCGCAGCAGCGC 480
Db 575 AAGGGGGCCCTTTGAGGCGCTCCGCTGCTGCTTCTACGTGATTAAGGCGCAGCAGCGC 634

```

QY	481	GAGACAGTACTGTCCATCTATAGAAAAGAGTACAGGAATGTGGCGACGTACAACTG	540
Db	635	GAGACAGTACTGTCCATCTATAGAAAAGAGTACAGGAATGTGGCGACGTACAACTG	694
QY	541	CTATCTGAATGCCCGCTTCATCTGCACAGATGTGGCAGTGGACTATGTTCCTAGCACC	600
Db	695	CTATCTGAATGCCCGCTTCATCTGCACAGATGTGGCAGTGGACTATGTTCCTAGCACC	754
QY	601	CTGTATCCGGAATGGCGCGGAGTACTATATCTCCCCACTGCTGCGCTCTGCGC	660
Db	755	CTGTATCCGGAATGGCGCGGAGTACTATATCTCCCCCAATGCTGCGCTCTGCGC	814
QY	661	CAATACTGCTGACCCCTGAAAAATCGGAGATTGCGCAAAACAGCTCTGTAAGTAA	720
Db	815	CAATACTGCTGACCCCTGAAAAATCGGAGATTGCGCAAAACAGCTCTGTAAGTAA	874
QY	721	GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACCTTTTACCGTCGAAATGCTGG	780
Db	875	GTTAACGATCGCTGTTTAAAGATCGGGTCGCGGCTTAACCTTTTACCGTCGAAATGCTGG	934
QY	781	ACAACAGACAGTATCAGACTGATTTCAAGCGCAACACCTTTATCCGATCCGACAGACC	840
Db	935	ACAACAGACAGTATCAGACTGATTTCAAGCGCAACACCTTTATCCGATCCGACAGACC	994
QY	841	AATACACGACACCGCGGACGACGATATTCGGGATACGAAGATATTCGACGCGCTGAAT	900
Db	995	AATACACGACACCGCGGACGACGATATTCGGGATACGAAGATATTCGACGCGCTGAAT	1054
QY	901	AATTGCTGAGAAAAAGAAATCCTAGCGCGCGCAGACCCCTCGTCCAGATAGCGTCCCGCAA	960
Db	1055	AATTGCTGAGAAAAAGAAATCCTAGCGCGCGCAGACCCCTCGTCCAGATAGCGTCCCGCAA	1114
QY	961	GAAATTCCTGCTGTAAACCAAGAAAGCGGAAGCGCGCACCCCGGACGCGCAAAAGCGGAA	1020
Db	1115	GAAATTCCTGCTGTAAACCAAGAAAGCGGAAGCGCGCACCCCGGACGCGCAAAAGCGTGA	1174
QY	1021	AAGAAGCGCCCTCCAGAAAGACTCGAGAGCAGACATGACAGGAGGCTTCTGAGAAAAAT	1080
Db	1175	AAGAAGCGCCCTCCAGAAAGACTCGAGAGCAGACATGACAGGAGGCTTCTGAGAAAAAT	1234
QY	1081	CCTGCCGCCCTCCCGAAGACGACGAGTCCCGGAGAGACACCGACGACGATGATCCAAAC	1140
Db	1235	CCTGCCGCCCTCCCGAAGACGACGAGTCCCGGAGAGACACCGACGACGATGATCCAAAC	1294
QY	1141	TCGATCCTGACTATTAACAATGACATGCCCCCGGTGATCCCGGTGAGAGACTACTAAA	1200
Db	1295	TCGATCCTGACTATTAACAATGACATGCCCCCGGTGATCCCGGTGAGAGACTACTAAA	1354
QY	1201	AGTTCTAATGCCCTCTCCATGCCCATATTCGGCGGTTCTGTAGCCTGCGCGCTCGCGCTC	1260
Db	1355	AGTTCTAATGCCCTCTCCATGCCCATATTCGGCGGTTCTGTAGCCTGCGCGCTCGCGCTC	1413
QY	1261	GTGGGCTACTGTTTGAGACATCGTAAATGCGCGCGTAG	1301
Db	1414	GTGGGCTACTGTTTGAGACATCGTAAATGCGCGCGTAG	1454
RESULT	5		
LOCUS	A22122	3502 bp	DNA
DEFINITION	ILTV gp60 gene.		linear
ACCESSION	A22122		PAT 22-JUN-1994
VERSION	A22122.1	GI:583833	
KEYWORDS			
SOURCE	Gallid herpesvirus 1.		
ORGANISM	Gallid herpesvirus 1		
REFERENCE	1		
AUTHORS	INFECTIONOUS LARYNGOTRACHEITIS VIRUS VACCINE		
TITLE	Patent: WO 9203554-A 51 05-MAR-1992;		
JOURNAL			

FEATURES	Location/Qualifiers
source	1..3502 /organism="Gallid herpesvirus 1" /db_xref="taxon:10386"
gene	339..3326 /gene="ILTV gp60 gene"
CDS	339..3326 /gene="ILTV gp60 gene" /codon_start=1 /protein_id="CAA01588.1" /db_xref="GI:583834" /translation="MGTMVLVHLFLVADALPTGRFCRWKVPFGTIQENLAVLA ESPVTGHATVPPEGAVSFQIFADPTLIRIRYGATEDELALERTSASDADNVTLS YRPREIHGAFITIGVATGOSTESSYVLSRLVNASLERSVRLERPCDENFLQNEP TRGSKRWLGPPSPYRDNDVAVLTAKOYIGECYSNSAAQTGLTSLNMTFFYSPKRI VTWTTGGPPSRITVYSSRENGOPVLKRVSDGFLVKYTPDIDGRAMINVIANSPADS GSVLAFTAFREGKLPVVIQLHRIDISGTEPEPGETTFDCOKMIETPYRALGSNVRD SIGGATLPPEFDIAPDFDGTSPPTTVPETATITLIPRSTSDMGFFSTARATGSET LSVPVQETDRLSTPLTLPLTGESENTLEPTAAGISTETPSAHEFTQTQSAEV VFTQSPSTESSETARSQSEPWYFTQTPSTEQALITQIATFVLFQTQPSDEOMTFT QTPGAEATEAPAQTPSTPELFTQSRSTPETARAPSAPEVFTQSSVTEVEFTQPS TVPKTTLSSSTEPALFTRTQSAQTEAFQTSSAGPDMRTQSTETHEFTQAPSTVPA TQTPSTEPEVLQSPSTPEPFIIRTLGAPEITQTPSTVEITQPSAPEVYTRSSS TMPETAOSTPLASQNPSSGTHTNTEPRVYQTPHTOKLYTENKTVSEPTVSEF HMSSTAESQTPLDVKEIVEKFSNDGEVATCVSTVSKPYRVETNMKVDLVDVMEIS GNSPAGVENSNEKQOLYRVTDGRTSVQIMCLSCSTKSPERYCLFDTSLIAREKDI APELYFTSDPQATYCTITLTPSGVVPREFEWSFNVSLEPYLATTVVSHTAGOSTVWKS SARAGEAWISGRGNIECTVLISDGT RVTTRKRCILNTWIAVENGAQAQAHVYSLFS GLVSLGCSISALYATLMTAIYE"
BASE COUNT	845 a 969 c 920 g 768 t
ORIGIN	
Query Match	27.7%; Score 361; DB 6; Length 3502;
Best Local Similarity	98.6%; Pred. No. 2.6e-84;
Matches	364; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	1 ATGCACCGTCTCATCTCAGACGGCGACCTGCGTTACTACGCCGAAAGAGAGGTGCTTAAC 60
Db	3134 ATGCACCGTCTCATCTCAGACGGCGACCTGCGTTACTACGCCGAAAGAGAGGTGCTTAAC 3193
QY	61 AAACACATGATTCGCGGTGAAAAACGGTCTGCTCAGGGCGCAGCTGTATTCTACTCTTTTC 120
Db	3194 GAACACATGATTCGCGGTGAAAAACGGTCTGCTCAGGGCGCAGCTGTATTCTACTCTTTTC 3253
QY	121 TGGACTTGTCTCAGGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
Db	3254 TGGACTTGTCTCAGGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 3313
QY	181 CATTTATTTTGAGGAATGCTTTTGACTATGCTACTGCTTCTTCTTCGCTAGCCAG 240
Db	3314 CATTTATTTTGAGGAATGCTTTTGACTATGCTACTGCTTCTTCTTCGCTAGCCAG 3373
QY	241 AGCACCGCCCGCTCACGTACGACTACATTTTAGGCCGTCCGCGCTCGACGCGCTAAC 300
Db	3374 AGCACCGCCCGCTCACGTACGACTACATTTTAGGCCGTCCGCGCTCGACGCGCTAAC 3433
QY	301 ATACCGCGGTTGGCCCGTATAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT 360
Db	3434 ATACCGCGGTTGGCCCGTATAACAGATACCTCACTAGGGTATCAAGAGGCTGCGAGCTT 3493
QY	361 GTCGAGCTC 369
Db	3494 GTCGAGCTC 3502
RESULT	6
LOCUS	AF081810 161046 bp DNA circular VRL 08-MAR-1999
DEFINITION	Lymantria dispar nucleopolyhedrovirus, complete genome.
ACCESSION	AF081810
VERSION	AF081810.1 GI:3822234
KEYWORDS	

SOURCE Lymantria dispar nucleopolyhedrovirus.
ORGANISM Lymantria dispar nucleopolyhedrovirus
Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
Nucleopolyhedrovirus.
REFERENCE 1 (bases 1 to 161046)
AUTHORS Kuzio,J., Pearson,M.N., Harwood,S.H., Funk,C.J., Evans,J.T.,
Slavicek,J.M. and Rohrmann,G.F.
TITLE Sequence and analysis of the genome of a baculovirus pathogenic for
Lymantria dispar
JOURNAL Virology 253 (1), 17-34 (1999)
MEDLINE 99124785
REFERENCE 2 (bases 1 to 161046)
AUTHORS Kuzio,J., Pearson,M.N., Harwood,S.H., Funk,C.J., Evans,J.T.,
Slavicek,J. and Rohrmann,G.F.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1998) Department of Microbiology, Oregon State
University, Corvallis, OR 97331-3804, USA
REFERENCE 3 (bases 1 to 161046)
AUTHORS Kuzio,J.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-1999) NCBI, 8600 Rockville Pike, Bethesda, MD
20894, USA
FEATURES
source Location/Qualifiers
1. 161046
/organism="Lymantria dispar nucleopolyhedrovirus"
/db_xref="taxon:10449"
39. 523
/function="enhancer; replication origin"
/rpt_family="hrl"
/rpt_type="dispersed"
673. 1410
/note="28772 Da; polyhedral matrix protein; Ld-PH;
Ldorf-1"
CDS
/codon_start=1
/product="polyhedrin"
/protein_id="AAC70186.1"
/db_xref="GI:3822235"
/translation="MHNFYNSPALGKTTYVDNKKYKNGITVIAKQAKRÖKHLËÖHEIE
ERSLDHLDRYLVAEDPFYGPCKQKLTLEKEIRNVKPDIMKLVNWSGKEFLRETWTR
FMEDSEPIVNDQEVMDIYLTINVRPTRPNRCYKFVAQHRLRDEGYVPHEVIRIVEPS
TYENNERYISLAKRGCGCPIRNHSAYTTSFEHFLNSVIMDNFYKPIVYVGTSAEEE
EILEVSLVEKIKEFADAPLFCGPAY"
complement(1418. 3085)
/note="61056 Da; Ldorf-2; Ld-Orf1629"
CDS
/codon_start=1
/product="essential structural protein pp78-81"
/protein_id="AAC70187.1"
/db_xref="GI:3822236"
/translation="MFKSYAGNEPESNLFHNEAYPPRDEALRLVSEVAAAPASACWL
IDKIMTSAVESPGVLTFDLREIYDILLRFENRVCAAAAPYRSDDSPRIRRLÖSMVDRLS
GHKALÖDLIARLYRTPDLSLEIKFLEAYRRHLVASRDERIDSIFNEIVELRDAPS
PSPSPERIFÖKSSTASAPAPAPAPARÖPAPAPARTPIAPADEPTROETPIAPASEP
ARÖESPISGSAPEPIRÖETPTGFLAPPPPPPPPPPEPELÖÖKSSAVPPPPPLP
PPGAPDDPGEIGÖEVRPKPAERAPTDLFAEIRGVÖLKPAATERAPTYTPDALFAE
IROGVKLPAAERADEPPKSSRAPLLIEIENRDKIKLKVPARATERPPASATNNPLM
ÖLNRKLESKMSSAESDANTSSWSDAEDDSLRLDALRIKLALLGPRLSSESEKR
IAKRLAGAKLKSAEKTLDELÖARAIEPDNPLSPYÖLTAPLYLHDLKLFESAVLDLF
NRGAYETALEKLEALÖVDIÖAPSLÖRMHDDISTYVKLÖKKRLESEA"
3087. 3911
/note="32438 Da; Ld-pk; Ldorf-3"
CDS
/codon_start=1
/product="protein kinase"
/protein_id="AAC70188.1"
/db_xref="GI:3822237"
/translation="MDALIGFADFHKECASARTALHLVNGKGVSVWKHGPTÖKSFF
YKRIEHKHFNAIEPFVHLMKFNKYFLRLFYSLHSLSREHLVMDYIPDGDFLMÖTE
PRLREPEISLIAYÖLIDALÖLHKHNVVNDVKLENVLYRREÖIYVCDYGLCKIAGS
PSTFEGTVDFSPÖKINKNAAVHDMWAVGVLLYEISTGKHÖPKLÖDESLDVETLH
KROIÖDVTFPADFNPFLEEFÖCFLGÖCYDYRAHSYEVIOÖNTYWKSIHVHWKÖR"
complement(4501. 7590)
/note="109196 Da; Ldorf-4"
CDS
/codon_start=1
/product="mucin-like protein"

/protein_id="AAC70189.1"
/db_xref="GI:3822238"
/translation="MCKNCAFHVITIKNLIAKSEVDFDSGSENEDRILVNRTCRTD
AÖPGPSRASDDKSDSDRDSKSGSDSEÖEPPSRDQÖPMLIPVEDKELNARR
KEILAKINEERAEAIKMLRDÖQAPPEYRPÖKSDSSSSSSSSSSSSGSDSDDECT
KNVSGKNIFÖNCYSRPEFEAKNVSEDLFEATGSEIAEAPAPAPADAEAPAPFAE
PADAEPAFDEESALÖEVPSPSSCLSLSPGPMÖFESVDEPATSGEHKDPÖISNSKA
STVSDKSRSPSADKHGSDRAKSPKSGCDRPPKSSAVSGKPKPEAKSKSSCDKPTGPG
VSTSGYDKPKSSADKPSDKAKSSSTVKPKSSADTPSDKAKSSSTVKPKSSSDKPSD
KAKSSSSSDKPCDKAKSSSTVKPKSSPDKPCDKAKSSSTVKPKSSSDKSPDKLSDKPS
DKAKSSSTVKPKSSPDKSPDKAKSSSTVKPKSSPDKSPDKAKSSSTVKPKSSPDKSP
KAKSSSTVKPKSSSAVPDPKPKHSDKSRDATVKKPKSSSAVPDPKPSDPSDGSRSFRP
PSESGRÖPKSLAÖTSPFLAALARDERLNRDLÖAAYASITGVREPEYPAPEPSA
PVDYFTLNSAEFAPEPAPEVPPEPSAPVDYFTLNSAEFAPEPAPEPAPEPAPEP
TSRFTSEPALGPVEPALEPEPALEPEPALGPGEPALEPEPALEPEGEPAEASEL
AVERPAÖPADLTSDAVESTFGHSTVASELAARPAELAPAPAPELTPEHSEPAÖ
PASKPASEPSAHLTPDRTLPSPLTVKSLÖSLMKSDDEVEEDTFVKNDEEDTVEED
IYIKNENNELCLVGPÖGYEHKASSSTPSASAEVNDVEVEIÖSEPHKAKPGVFMMN
KRKKSÖLMSVRRKRPYSÖQPÖPPPPPPPPPPPPPSFASPYASASAHÖYPPMYHMSPY
HLPÖPSFVPVPPÖQFPÖRYVYAPÖTALPDHCYAPAEVPDHÖEVLDDILÖTCLGPDÖTETPÖ
7722. 8291
/note="21763 Da; Ldorf-5"
CDS
/codon_start=1
/product="Ldorf-5 peptide"
/protein_id="AAC70190.1"
/db_xref="GI:3822239"
/translation="MRTLVPHDRHILFDVAVDLHDRVÖGVVRIIHAYFLGHSVFI
ÖLEHAFÖVILTLDHHTVYVFLGLESLEGYDLEFVIFYVLFYÖKLDVLOHÖAHÖLÖI
VLÖTLHLHSHTVLHLVLEHLVGHVPAESVLKEVDPYKLEVARVÖNGHHVKLLALVÖG
NVVYLFYLLIVEADLLPGLIARSVYT"
complement(8282. 8524)
/note="8960 Da; Ldorf-6"
CDS
/codon_start=1
/product="Ldorf-6 peptide"
/protein_id="AAC70191.1"
/db_xref="GI:3822240"
/translation="MSEPDGRGRARGAÖQTYRADGAARRLIARLLIVSIGRLIASRT
RSGLDGSIÖRYRGK LISRTVRYVRAIVLÖKÖFLTÖQV"
8619. 9194
/note="22146 Da; Ld-g22; Ldorf-7"
CDS
/codon_start=1
/product="Ldorf-7 peptide"
/protein_id="AAC70192.1"
/db_xref="GI:3822241"
/translation="MSLÖRTINHAÖVLLTRFDGFSVÖTKVNPCLFMFKDKNFRDÖV
GÖYVRFÖSKVSDSKTKVPKGYVLCÖKYOVCVHDKLYGTRDPLKRENVVPPNPGNVIVÖ
YMLVREDCVDACRLEKÖDARPENÖYVLYGSEFPREHLÖLSMTLDTKYCSYÖDWSGÖTT
MSKLNAAÖRDLDKSLAGSVNÖKTSNIÖVCN"
9496. 10019
/note="10019 Da; Ld-ö22; Ldorf-8"
CDS
/codon_start=1
/product="Ldorf-8 peptide"
/protein_id="AAC70193.1"
/db_xref="GI:3822242"
/translation="MISSVÖQTUSASPRLSVTFAEKÖLRSKNNFYILFVSÖKNKLI
ALIVCYLFEKÖLYLHPVRIAEKÖAYRVDLLRPVÖKTSLSNRALAOADRÖRRPTFÖALÖ
KSGLRIÖDPYÖTEÖTLRS"
10567. 11085
/note="19624 Da; Ldorf-9"
CDS
/codon_start=1
/product="Ldorf-9 peptide"
/protein_id="AAC70194.1"
/db_xref="GI:3822243"
/translation="MSLSÖQTEAARADKRLKÖEDCSGVAKKLCILEAAÖQAEAEPRP
TMWLYKDADADIEYSDÖEYEFDEMEEGÖWDRDLIFÖTDPDÖPSAAPGÖGÖTRICK
YÖTMFAGÖTHRKGDWÖKÖFÖKÖLÖPLDLKIOGLÖQÖTVDAVYGSRKKEVEDVÖTLWLKDD
DRVFLIVSVLLI"
complement(11115. 11336)
/note="8543 Da; Ldorf-10; Ldorf-10"

CDS
/codon_start=1
/product="Ldorf-10 peptide"
/protein_id="AAC70195.1"
/db_xref="GI:3822244"
/translation="MILADWGA YOSITSKTKHCNLSNPRYRLTSKTKHCNLDSSSYQ
RLTSKTKHCNLSNKPILYPIIILHYKSFK"
11530..12264
/note="27768 Da; Ldorf-11; Ldorf-11"
/codon_start=1
/product="Ldorf-11 peptide"
/protein_id="AAC70196.1"
/db_xref="GI:3822245"
/translation="MSKRIC SASDSESDRPRGLYGDALIEALGLPKESDRPGKRPRL
EILSKNVHGDKLDVYKAFGAGNSELESDDTLAMSSKRPEGSELESETDIDLAM
PSGADYEKYSKSLIKLITKTLVYLRKLDKLLIKQKLAIO LKNKIYSEALRKE
FAELQCHFRNNYSVRPILMLPTLEQLRAIRDWCIA TSAKIDAFIEDPVMRNGNSNI
HCFEIDVFIDFCVYMOKAENLFLKD"
12647..13165
/note="20383 Da; Ldorf-12; Ldorf-12"
/codon_start=1
/product="Ldorf-12 peptide"
/protein_id="AAC70197.1"
/db_xref="GI:3822246"
/translation="MELYNDNIKKECLIKLIIIDFVFRYLNLNQIEKEFEVFKOKN
INSLVDTLMAHYATETLHNSRFEEKRNRIEONRQALLRRIEONRQALOKOR
EHSKRKRQAEAPLQLKEDFPELAEEDEFETLDAEAAA AVSKDESPOGESPOVESI
OEGAKMRUDEA"
complement(13282..13482)
/note="7643 Da; Ldorf-13; Ldorf-13"
/codon_start=1
/product="Ldorf-13 peptide"
/protein_id="AAC70198.1"
/db_xref="GI:3822247"
/translation="MILDILYFIKSGYISYTLAVYLFQCLIFLYLIKSGYISYTLAAK
CAIGLIVGFHKELKONYITTY"
13685..14755
/note="38636 Da; occluded; Ldorf-14; Ld-odv-e56"
/codon_start=1
/product="envelope protein"
/protein_id="AAC70199.1"
/db_xref="GI:3822248"
/translation="MSFESSLRVNKVYPNPNOFLAVDNLRYLNSPENGQNVESAPS
TRQIGTNRFPVGYNLFPNRFVPSDDINRYMRNDTAGMRNIFTSADNSQLNSLGQLR
VDNIPDAGLHAFLRQGVKSNYPSTNTRPGVEGLATNPRHLRNLGLKSVGTAL
LIGGVYLVFSATLVQDLVSLNRVGSYYVGSNGDSSVCLLHRTCOLDTSL
SDVSVCOFDPLPDPPAGORAICNGFNERYEOSVCRASDPYADPSEQYVIDSLPAG
HTLMCLEPYNLGLDGLDHLGEEGLVDKSSNSKDLSSKMLPVLVALLIGIFAVV
LILFEVFKRLMNKOTVIDIRANA"
complement(15221..16921)

CDS
complement(13282..13482)
/note="7643 Da; Ldorf-13; Ldorf-13"
/codon_start=1
/product="Ldorf-13 peptide"
/protein_id="AAC70198.1"
/db_xref="GI:3822247"
/translation="MILDILYFIKSGYISYTLAVYLFQCLIFLYLIKSGYISYTLAAK
CAIGLIVGFHKELKONYITTY"
13685..14755
/note="38636 Da; occluded; Ldorf-14; Ld-odv-e56"
/codon_start=1
/product="envelope protein"
/protein_id="AAC70199.1"
/db_xref="GI:3822248"
/translation="MSFESSLRVNKVYPNPNOFLAVDNLRYLNSPENGQNVESAPS
TRQIGTNRFPVGYNLFPNRFVPSDDINRYMRNDTAGMRNIFTSADNSQLNSLGQLR
VDNIPDAGLHAFLRQGVKSNYPSTNTRPGVEGLATNPRHLRNLGLKSVGTAL
LIGGVYLVFSATLVQDLVSLNRVGSYYVGSNGDSSVCLLHRTCOLDTSL
SDVSVCOFDPLPDPPAGORAICNGFNERYEOSVCRASDPYADPSEQYVIDSLPAG
HTLMCLEPYNLGLDGLDHLGEEGLVDKSSNSKDLSSKMLPVLVALLIGIFAVV
LILFEVFKRLMNKOTVIDIRANA"
complement(15221..16921)

Query-Match 4.0%; Score 52.2; DB 14; Length 161046;
Best Local Similarity 50.2%; Pred. No. 0.012;
Matches 129; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
CDS
999 CCCGGACGCAGAAAGCAGGAGCCCTCCAGAGAAGACTCGAGAGACACATGCA 1058
Db 87810 CCGCGACGAGACAAATCGAGGTGCGGACGAGACGAGACGACAAATCGA 87869
QY 1059 GCGAGAGGCTTCTGGAGAAATCCTGCCGCTCCCGGAAAGACGACGAAGTCCCCGAGGA 1118
Db 87870 CGAGGACGAGACGAGCAAAATCGACGAGACGAGACGAGACGAGACGAGCA 87929
QY 1119 CACCGACGACGATGATCAAACTCGATCCTGACTATTTACATGACATGCCCGCGGTGAT 1178
Db 87930 GGACGAGGACGAGACGAGACGAGACCAAGAAATGTCGTTCCACGCCGCCGACCTCAC 87989
QY 1179 CCCGGTGAGAGACTAATAAGTTCTAATGCCGCTCTCCATGCCCATATTGCGGCGGTT 1238
Db 87990 GCTGAGCGATAGATGATTGATTTTTCGCTCTTCTACTCGGTGCCCAAGTTGAAACGAT 88049
QY 1239 CGTAGCCTGCGCGGTGC 1255
Db 88050 CGCGTGCTCCGCCGCCG 88066

RESULT 7
MXARPOD 2676 bp DNA linear BCT 26-APR-1993
LOCUS
DEFINITION M.xanthus sigma factor (rpoD) gene, complete cds.
ACCESSION M32347
VERSION M32347.1 GI:150117
KEYWORDS sigma factor.
SOURCE M.xanthus (strain FB (D2F1) DNA, clones pSIG01,2).
ORGANISM Myxococcus xanthus
Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
Myxococcales; Cystobacteriineae; Myxococcaceae; Myxococcus.
REFERENCE 1 (bases 1 to 2676)
AUTHORS Inouye,S.
TITLE Cloning and DNA sequence of the gene coding for the major sigma
factor from Myxococcus xanthus
JOURNAL J. Bacteriol. 172, 80-85 (1990)
MEDLINE 90094274
FEATURES
source
location/Qualifiers
1..2676
/organism="Myxococcus xanthus"
/db_xref="taxon:34"
298..2424
/note="sigma factor (rpoD)"
/codon_start=1
/transl_table=11
/protein_id="AAA25404.1"
/db_xref="GI:150118"
/translation="MPTQKPSKASVKTKKKVDPVIRKKKAPDGAAGAKTGAEEKE
ELVARDVVAEATEKLKKKKGVAPAMGDDVDPEEAEEAAA AVQVDPDAVEDIEEDP
VAERKEVKDLLAAGREKGFITYDEVNDLPADIVSSDQIDVMSMFGNDIEIVDAQ
AAQNNELKPTVTVEEEKEDADEDEDEDEDEPGKSNPVRLLYLRKMSVSLITREGEV
EIAKRJEDGEKEVLRALLACVAVEEILDIGNKLTAKLRVRDVIKDAPEETQSEGAE
EAPEEVGEEDAPAOQAQSELNKIEQIKQIERFRKRAQDCDVLBEELSKKLTVEVK
KEVQEIKDRLTKMMEVLEENLNKKQVDRIVINLAKLIERVDKAEDELRLERYDC
SMKELRPOLKESRENPNIGKLOKOLNLTPEQLEVLDRDVRTAVRKIKVEEENLPV
ESLRNVDAIRLGERRAEAKSELVEANLRVVSIAKKYTNRGLQFLDLIOEGNIGIM
KAVDKFEYKRGYKFSVATWMIROAITRAIADQARTIRIPVMIETINKLIRTSRYLV
QETGREPTPEEIAEKMEIPLDKVRKVLKIAKEPISLETPIGEEEDSHGDFIEDKSLV
SPADAVINMNLAEQTRKYLATLTPREEKVLRMRFIGIGEKSDHTLEEYGDFFEVTRERI
ROIETAKLRKLRLHPSRKRIRSFVES"
2451..2478
/note="pot. transcription termination signal; putative"

CDS
/organism="Myxococcus xanthus"
/db_xref="taxon:34"
298..2424
/note="sigma factor (rpoD)"
/codon_start=1
/transl_table=11
/protein_id="AAA25404.1"
/db_xref="GI:150118"
/translation="MPTQKPSKASVKTKKKVDPVIRKKKAPDGAAGAKTGAEEKE
ELVARDVVAEATEKLKKKKGVAPAMGDDVDPEEAEEAAA AVQVDPDAVEDIEEDP
VAERKEVKDLLAAGREKGFITYDEVNDLPADIVSSDQIDVMSMFGNDIEIVDAQ
AAQNNELKPTVTVEEEKEDADEDEDEDEPGKSNPVRLLYLRKMSVSLITREGEV
EIAKRJEDGEKEVLRALLACVAVEEILDIGNKLTAKLRVRDVIKDAPEETQSEGAE
EAPEEVGEEDAPAOQAQSELNKIEQIKQIERFRKRAQDCDVLBEELSKKLTVEVK
KEVQEIKDRLTKMMEVLEENLNKKQVDRIVINLAKLIERVDKAEDELRLERYDC
SMKELRPOLKESRENPNIGKLOKOLNLTPEQLEVLDRDVRTAVRKIKVEEENLPV
ESLRNVDAIRLGERRAEAKSELVEANLRVVSIAKKYTNRGLQFLDLIOEGNIGIM
KAVDKFEYKRGYKFSVATWMIROAITRAIADQARTIRIPVMIETINKLIRTSRYLV
QETGREPTPEEIAEKMEIPLDKVRKVLKIAKEPISLETPIGEEEDSHGDFIEDKSLV
SPADAVINMNLAEQTRKYLATLTPREEKVLRMRFIGIGEKSDHTLEEYGDFFEVTRERI
ROIETAKLRKLRLHPSRKRIRSFVES"
2451..2478
/note="pot. transcription termination signal; putative"

Query Match 3.5%; Score 45.8; DB 1; Length 2676;
Best Local Similarity 57.2%; Pred. No. 0.4;
Matches 83; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 1045 GAGACGACATGCAGCAGAGGCTTCTGGAGAAATCTGCCGCCCTCCCGGAGACGAC 1104
Db 511 GATGACGTGGACCCGAGAGAGGCCGCGGAAGAGGCCGCCGCCGCTCCAGGTGCAACCCC 570
QY 1105 GAAGTCCCGAGAGACACCGACGATGATCCAAACTCGGATCCTGACTATTACAATGAC 1164
Db 571 GACGCCGTGAGAGACGACATGAGAGAGATCCCGGTGCGCCGCAAGGAGGTCAAGGAC 630
QY 1165 ATGCCCGCGGTGATCCCGGTGAGG 1189
Db 631 CTGCTCGCCGCGGCGCGGTGAGAGG 655
RESULT 8
MXU20669 7980 bp DNA linear BCT 21-MAR-1997
LOCUS MXU20669
DEFINITION Myxococcus xanthus ribosomal protein S21 (rpsU), unknown protein,
primase (dnaG), major sigma factor (rpoD), and putative
methylguanine-DNA methyltransferase genes, complete cds.
ACCESSION U20669
VERSION U20669.1 GI:710339

KEYWORDS
SOURCE Myxococcus xanthus.
ORGANISM Myxococcus xanthus
Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
Myxococcales; Cystobacteriinae; Myxococcaceae; Myxococcus.
REFERENCE
AUTHORS 1 (bases 4593 to 7280)
TITLE Inouye,S.
JOURNAL Cloning and DNA sequence of the gene coding for the major sigma
MEDLINE J. Bacteriol. 172 (1), 80-85 (1990)
90094274
REFERENCE 2 (bases 1 to 7980)
AUTHORS Davis,J.M., Mayor,J. and Plamann,L.
TITLE A missense mutation in rpoD results in an A-signalling defect in
Myxococcus xanthus
JOURNAL Mol. Microbiol. 18 (5), 943-952 (1995)
MEDLINE 96422481
REFERENCE 3 (bases 1 to 7980)
AUTHORS Plamann,L.S.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1995) Lynda S Plamann, Department of Biology,
Texas A & M University, College Station, TX 77843-3258, USA
FEATURES
source 1. .7980
/organism="Myxococcus xanthus"
/strain="DK5060"
/db_xref="taxon:34"
misc_feature 1. .7980
/note="rpsU is the first gene in the rpsU-dnaG-rpoD
macromolecular synthesis operon"
1887. .2081
gene /gene="rpsU"
1887. .2081
CDS /gene="rpsU"
/note="similar to E. coli ribosomal protein S21: SwissProt
Accession Number P02379"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ribosomal protein S21"
/protein_id="AAB60205.1"
/db_xref="GI:710340"
/translation="MPGIRVKEGESIESALKRFFKATEKAGILSEIRKREHYEKPVS
KKKALAAKKRAVKKARKSF"
2117. .2644
CDS /codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAB60206.1"
/db_xref="GI:710341"
/translation="MSGALAPAVEYDLNNVLPSSGAATMATLKERIDADLDKAMRSKNE
LTLVLRMLKSAVKYKEVEPGASALDDAGVLTITAGLIKQRDSVEQFKTGRPELAE
KEEAETISILQNVLPKQLTADELAERAAIAETGAKPKDMGAVMKNLNAKIHGKAEG
RAISEAVKSELAKLS"
2977. .4797
gene /gene="dnaG"
2977. .4797
CDS /gene="dnaG"
/note="similar to E. coli primase: SwissProt Accession
Number P02923"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="primase"
/protein_id="AAB60207.1"
/db_xref="GI:710342"
/translation="MGVVIPEHKIQEVLERVDLVGLISRHAVDLKKAGREWKACCPFHQ
EKTPSFYVPEKRFYFCHGCRASGDAVSFQRYLGKTFDLDAVDLARELGVDLEAEQD
PSMREKQIKEATDQAAEHFRAMLMOQDEGSRARAYIASRGVDETAMAFGLGWAPLE
WASLTERFQKLGMLEWAAKAGLVLRKNSGDGYDDFFRSRVMPVIRAEGRPIAFGGRLL
IGADEGPKYLNRESRLYNKSETEFGMDQSRDEIRKKAAYLVEGYFDALGLHGVGR
HAAVLCSTNLTAGHMOVLKRAEARELLILLDGDAGLAVERISGPLLAAGATARVAL
LPGDDPDTFARREGQEGVERLEGAHPLTSHLFASLLPEKKAASFEEKMAALERLKP

VGQVPVGLVTRATLESAAVEHFGWRPADVEAALRSKVLPLPKPAGDAPSSPNRPAPP
LEKPPALCEFYGAVLKEPRLMARDTFVCDLSHGLMALAHATSGHGANDALE
SSEAVKRGIESALRQLPSEPVPLAEAFSLICREIMVRIDERLYVIKRATEQTPGAFD
LTEETRQLIVERVELLALKKRVLEELKPASSGTTAPMQPV"
4890. .7016
gene /gene="rpoD"
4890. .7016
CDS /gene="rpoD"
/note="similar to E. coli sigma factor: SwissProt
Accession Number P00579"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="major sigma factor"
/protein_id="AAB60208.1"
/db_xref="GI:710343"
/translation="MPTQKPSKASVKPTKKKVDPVIRKKKAPDGAAGAAGAKTGAEKE
ELVARDVVAETEKLLKKKGAVAPAMDVDVPEEAAEEAAAVQVDPDAVEDDIEDP
VAERKEVKDLLAAGREKGFITYDEVNDALPADIVSSDQIDDVMSMGDNDIEIVDAQ
AAQNNIEIKPTVVEEEKEDADEDEDEPGKSDNPVRLYLRKGSVSLTREGF
EIAKRIEDGEKVLRALACKVAVEEIDIGNKLTAKLVRVDYKDAPEETQSEGA
EAPEEVGEQDAPQAQSELNKIEQIQIERFRKFAQDCDVLLEELSSKKLTREVRC
KEVKQEIKDRLTKMVEVLEEMRLNKQYDRIVINLKALIERVDKAEDELRLDERYDC
SMKELRPOLKRESRENPNIGKLOQLNLTPEOLEVLDRDVRTAVRKIKVEEENLPV
ESLRNNDAIRLGERRAERAKSELVEANLRLVSIAKKTYTNRGLOFLDLIOGNIGLM
KAVDKFEYKRGYKESTYATWIRQATRAIADQARTIRIPVHMETINKLIRTSRLV
OEIGREPTPEETAEKMEPLDKVRKVLKIAKEPISLETPIGEEESHLDGFIEDKSLV
SPADAVINMNLAEOTRKVLATLTPREKVLRMRFIGEGKSDHTLEEYGGDFEVTERRI
RQIEAKALRLRHPRSRSLRSFVES"
6680. .6681
variation /gene="asgC767"
/note="a mutant allele of rpoD; the asgC767 (rpoD-EK598)
mutation results in a glutamate to lysine substitution"
/replace="AA"
6680. .6681
gene /gene="asgC767"
7040. .7483
CDS /note="putative methylguanine-DNA methyltransferase;
similar to E. coli methyltransferase: SwissProt Accession
Number P09168"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/protein_id="AAB60209.1"
/db_xref="GI:710344"
/translation="MGPRSIRGPGAFVPLESGAVALRAKTKTPKAEVKKAPLPFSKAV
WKAVRAIPRGQVRSYAQVALYAGRPGAAGVGREMATLPQOPELPLPWWRVTRSDGTL
APQVAQEQARRLRAEGVEVTQRGDTFRVKVPREVPAAGLATKRQR"
7049. .7051
conflict /citation=[1]
/replace="G"
7131 /replace="G"
/citation=[1]
/replace="CG"
7165. .7185
conflict /citation=[1]
/replace="GCGTTGAAGCGTCC"
7199 /citation=[1]
/replace="G"
7207 /replace="G"
conflict /citation=[1]
/replace="G"
7222 /citation=[1]
/replace="G"
7251 /replace="G"
conflict /citation=[1]
/replace="G"
BASE COUNT 1465 a 2732 c 2660 g 1123 t
ORIGIN

exon	gene	CDS	intron	exon	intron	exon	intron	exon	gene	CDS	intron	exon	intron	exon	gene	CDS	intron	exon	repeat_region
<pre>/gene="B1IN2.020" complement(12353. .12531) /gene="B1IN2.020" /number=1 complement(join(12353. .12531,12616. .12798,12825. .12981)) /gene="B1IN2.020" complement(join(12353. .12531,12616. .12798,12825. .12981)) /gene="B1IN2.020" /codon_start=1 /product="hypothetical protein" /protein_id="CAD11396.1" /db_xref="GI:16944632" /translation="MLPDGDLALWRSLLVLAGSRDSSAHFPTHTRHPQASLIOEY GTGAVIVVHSDHYTPARSWSPVRHWPVPAPPEFAFWSPKAPRAKHVVDHPKGGDG EAGIPSEPSWRDVGTVPEATEPSQRQRAAGEQPAGAE DNSQESTQDSTRPSDRVOLEMS GPNRVKRIIVEG" complement(12532. .12615) /gene="B1IN2.020" /number=1 complement(12616. .12798) /gene="B1IN2.020" /number=2 complement(12799. .12824) /gene="B1IN2.020" /number=2 complement(12825. .12981) /gene="B1IN2.020" /number=3 complement(join(14115. .16258,16397. .16536,16654. .16958)) /gene="B1IN2.030" complement(14115. .16258) /gene="B1IN2.030" /number=1 14115. .16958 /gene="B1IN2.030" complement(join(14115. .16258,16397. .16536,16654. .16958)) /gene="B1IN2.030" /note="strong similarity to CUTINASE TRANSCRIPTION FACTOR 1 BETA, Fusarium solani, SWISSPROT:CTIB_FUSSO" /codon_start=1 /product="probable CUTINASE TRANSCRIPTION FACTOR 1 BETA" /protein_id="CAC28684.1" /db_xref="GI:12718376" /db_xref="SPTREMBL:Q9C2C5" /translation="MSDSGEGSQCTPSPSPSTENESTGVPCTSPCEEKPPSLARTEAA TPSKRRPSKALPADSONPRPAKRRAACLTGRNRKYRCNVVEGLPCNNCKWGGRM KNMLAGQAVPNGVAAYGGEFGARTLGHQPIILSSQOLAHANASLMADAQATOTQGLMN NPSIPETSARLMQNPQLMTMPGNLTNP SATAVNDRPTISVSTPSQPTLQPNLQPK PNRTVSPNLPNFFKPLPPKDAVDVQYMAKGAFTIPTEVQNALMKAYIEYHPYML LELRQFLTAHSNGSQISLVLYQAVMFAGSNFYAQKYLDAGLGSRRNARKELFMR TRVLVDCDVEKDRLDLVQALLMTYWEETPEDQDTHWIGVAISLGLTGIHRNPAN LAMPRAQKALMKRIWCCCFMRDRILALGMRRPTRKDEDFVPMLEESDFEIVELPAD NQLGPNCAVVRNATOROLARLCTIOKARICVAISHMIKTQYTVLNHGGGLPAGQTS GTMLEFPNKSLLNIQEVQKVDOMLESWRLQLPEDCOYRPLLTEALAEEDQPVAVHRTL LHMVYHTTVSALHRPHYLTMEQAATQPAQTSILAQQARSKVHAATMVTMAEDLLRH GLAKYLPITTAVTITLPAMTVHLHSRSPDELSQARRDFEVCAKLLQLRGMTAAAE FAHGFGLMGVEARHKATAVSPQGPQAPVSLGHQVHLGPEKQETPPPETTSFSPSTNTT FYRPGQADAMVSGVGHGALVLVNVNDGGSTPQTQDVESSAGLTPVDTFTAYEEV QTAMECDITNGAMANYFDDAFGSILDPPEAVEGDEMAFTMSHENDDEWMEFDGPCNSNV VPAAYVC" complement(16259. .16396) /gene="B1IN2.030" /number=1 complement(16397. .16536) /gene="B1IN2.030" /number=2 complement(16537. .16653) /gene="B1IN2.030" /number=2 complement(16654. .16958) /gene="B1IN2.030" /number=3 22394. .22415</pre>																			

	repeat_region	/note="22 bp poly(t) tandem repeat" 23093..23118 /note="26 bp poly(a) tandem repeat" 23601..23682 /note="82 bp aagaagaagaagaagaagaagaagaagaagaag agaagaagaaga... tandem repeat" complement(24165..26144) /gene="B1IN2.040" /number=1
	exon	24165..27082 /gene="B1IN2.040" complement(join(24165..26144,26234..26262,26290..26717, 26829..27082)) /gene="B1IN2.040"
	gene	complement(join(24165..26144,26234..26262,26290..26717, 26829..27082)) /gene="B1IN2.040"
	CDS	complement(join(24165..26144,26234..26262,26290..26717, 26829..27082)) /gene="B1IN2.040" /note="strong similarity to beta-glucosidase (EC 3.2.1.21) 1 precursor, Aspergillus aculeatus, PIR:J04939 contains ATP/GTP-binding site motif A (P-loop) [ALGANGS]:glycosyl hydrolases family 3 active site [LLKDELGFQFVMSDWLS];Protein kinases signatures and profile [LGMGWSSGVANYPFDPPITAIK]" /codon_start=1 /product="probable beta-glucosidase 1 precursor" /protein_id="CAC28685.1" /db_xref="GI:12718377" /db_xref="SPTRMBL:O9C2C4" /translation="MAVRSNFWLTALSFDDLTCNASPVQRRAVPDGYAAPYYPT PYGWEWSMKDSYAKAOLVGKMTIAEKNITGSGMFMGNPCVNGSAVRVGPOL CLDGALGVNTDHNTAFAPAGITTATFDKDLIYARAVALKEFRGAHVLGPMVG PLGRKPLGGNWEFGADPVLQIGALTIKGVQEQIVATIKHLIGNOEVMRYNP FOAGYSNNIVYSRLTQLILMPFADGVHAGVGAAMTSYNNAVNGSAASONAYLINLLK DELGFGFYMSDWLSHISGSVASALAGLMSMGDTQIPLFNSPFKEHLTEAVLNGSV PVDRLNDMATRIVAAMYQFGQDKNEPAVFHSVSSEKGLLYPALPVPVIGKVMFV DVQADHGVSAROVAQDAITLLKNDNFPLSTKSSLRIEGSDARVDPGPNACGNRAC NKGTLMGWSGVANYPFDPPITAIKKRYENVKL YDSDDFPHTLFPSPDDDAIVF INSAAGENSLTVEGNHGDNDKFSAMWNGDNLVOKAENYKNVIYVHTVGPLLEP WIDLPSYKAVLFAHLPGQAGESLANVLEGDSPSHLPYSTIKKENLPSVTKLVK EIIGQPDDTYSEGLEYIDYRWLNKGQIKPRYAFGHLSYTTFNTHAHKIIVNALSSAL PPARQKPSPVAAILSTEIRPASDAYEPAGSKVWRXIIYSWLSDADAKRAYAVGTSSSSK SGSQITYPYPEGYSVOKPVPAGGOGGNPALEFTILELDVTQONTGSRHKASVOA YIQPTDSGYDTPIIQLRDFAKTELGESETVTLRLRRKDL SVMDTQKONWVAPGA LGANGSKSRYIWLGEESDKLFTRCFSDPLVCERGEVPEPV"
	Intron	complement(26145..26233) /gene="B1IN2.040" /number=1
	exon	complement(26234..26262) /gene="B1IN2.040" /number=2
	Intron	complement(26263..26289) /gene="B1IN2.040" /number=2
	exon	complement(26290..26717)
	Query Match	3.3%; Score 43.6; DB 8; Length 99360;
	Best Local Similarity	56.2%; Pred. No. 2.1;
	Matches	82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
OY	988 GAAGGGCGACCCCGAGCGAGAAGAGCAAGCAAGAGGCCCTCCAGAAACTCGAG	1047
Db	51436 GATGGGCTCTCGCGCCCTCGATCTCGCGAGCTGCTTGCCGGTGTGAGCTCTGTAG	51495
OY	1048 GACGACATGCAGGCAGAGGCTTCTGGAGAAAATCTGCGGCCCTCCCGGAGAGCAGCAA	1107
Db	51496 GACGGCCTGCTCCGCCGAGGGTGTGTCGCGAGGATCCTACGCCCATTTGTGAGGGCTACGGA	51555
OY	1108 GTCCCCGAGGACACCAGCAGCATGA	1133
Db	51556 ACCTCCTGTGAATAATGGGCTGGATGA	51581

RESULT 11

AF429315/c	125020 bp	DNA	linear	PRI 18-JAN-2002
LOCUS	AF429315			
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.			
ACCESSION	AF429315			
VERSION	AF429315.1	GI:17646244		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 125020) Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L.			
TITLE	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2			
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)			
MEDLINE	21583737			
PUBMED	11694876			
REFERENCE	2 (bases 1 to 125020)			
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA			
FEATURES	Location/Qualifiers			
SOURCE	1. 125020 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q24.3; between D16S520 and WI-12410" /note="Isolated from a patient with Huntington's Disease-Like 2 (HDL2)" complement(35581..35746) /rpt_type=tandem /rpt_unit=ctg complement(<36507..>36887) /gene="JPH3" /product="junctophilin 3" complement(<36507..>36887) /gene="JPH3" /note="Jp3" complement(<36507..36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctophilin 3" /protein_id="AAL40941.1" /db_xref="GI:17646245" /translation="MSSGRFNFDDGSGYGGWEDGKAHGHCVTGPKGGEYTSWSHGFEVLGVYTWPSGNTYOGTWAQGRHGIGLESGKWKVYKGEWTHGFKRGYGVREACANGAKYEGLTWSNGLQDGYGTETYSDG"			
CDS				
mRNA				
gene				
repeat_region				
BASE COUNT	29056 a 32731 c 30696 g 28283 t 4254 others			
ORIGIN				
Query Match	3.3%; Score 43.4; DB 9; Length 125020;			
Best Local Similarity	12.6%; Pred. No. 2.5;			
Matches 117; Conservative 382; Mismatches 409; Indels 18; Gaps 4;				
QY	217 CTGCTTCTCTTCCTTCGCTAGCCAGACGACCGCGCGCTCAGTACGACTACATTTTAGGC	276		
Db	52077 YKGSWGSYWASMSRSYKSCASCGWSTSYRSKRGGCWSATSYKCMYGSKTRSSGTGKRSAC	52018		
QY	277 CGTCGGCGCGCTGCAGCGCGCTAACCATACCGCGCGGTGGCCCGGTATAACAGATACCTCACT	336		
Db	52017 MCMKWCWSSNSRSMCMSTAYMKCWGKMYWYRYSRKRYSCTYRGSCKYKSRMAKSMMRGS	51958		
QY	337 AGGATATCAAGAGCGCTGCAGCGTGTGCAGCTCAACCCGATTTCTAAGCTGGACGACATG	396		
Db	51957 MRSSY--YWMRMCMWKGRCRRGAGSAARSYKKGAMKSGMRGCGSTYYGKMGCGRSKS	51900		
QY	397 ATATCGGCGGCAAGAAAAAGAAAGGGGGGCCCTTTGAGGCGCTCCGTCGTGGTTC	456		

```

Db 51899 MYTGSMKCAATGKMSWSMCMGRCAAMWSKGSYYRGWGCWCMASYWMCASYSYSRMRGSAMS 51840
      :: : | : : : | : | : | : : : | : : : | : : : | : : :
QY 457 TAAGTGAATTAAGGGCGACGACGGCGAGGAGACAAGTACTGTCCAACTCTATAGAAAAGAGTAC 516
      :: | : : | : : : | : : : : : : : : : : : : : : : | : : : | :
Db 51839 YWAGGYTSYRKGSWGWGKMYKGSARAGSMSRWSKRSWWMKSKGRSARASARWKSYSACAS 51780
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 517 AGGGAATGTGGCGACGTACAACTGCTATCTGAATGCGCCGTCATCTGCAC - - - - AGA 571
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51779 RKKSCWRMRMRCGSYMTGRSSMGTMKRRKGSRTSMGSKAGTYRSARRAAMYRSCSKGRRRG 51720
QY 572 TGTGGGCGAGTGGACTATGTTCTCTAGCACCCCTTGATGCGGAAAATGGCGGGAGCTACTA 631
      : : : : : | | | : : : : : : : : : : : : : : : : : : |
Db 51719 KKKSSKGGGRGGKMMASMTYSKRCCAKSTGRRRGDTKGSSCWTKRKGWGRTRKAAAWCA 51660
QY 632 TAATCTCCCCCACTGTGCGCTCTCTG - - - - - GCCAACTCTGCTGACCCCTGAAA 681
      : : : : | : | : | : : : : : : : : : : : : : : : : : :
Db 51659 KCCWMMYRCTGCMCAKTYAATYKWSKSRWTTGWRMGTTRAMWMMWYWMKKYRARMKC 51600
QY 682 ATCGGGAGATTTGGCGCAACAGCTCTCGTAAGTCTAGAAGTAAAGATCGCTGTTTAAAG 741
      | : : : : | : : : : | : : : : : : : : : : : : : : : : : :
Db 51599 AKTRSMWMAASKYSKSYARKCCSMMASAYKKYKSRASRCKGWGKNAKSCMRRYMMYKWWGGA 51540
QY 742 ATCGGGTCCGAGCTTAACCTTTTAAACCGTCGAATGCTGACACAGAACAGTATCAGACT 801
      : | : : : : : : : : : : : : : : : | : : : : | : : :
Db 51539 GKCCSAGGMTGRTGSRWSACCMKMSKKSASKARYAYASTGAAMSMCGTCYMKMCTRRMM 51480
QY 802 GGATTTCAAGGCGAACCCTTTATCCGATCCGACAGACACCAATACAGACACGCGGACGAC 861
Db 51479 ATACAMRRWRAMWMMSCMRKCTSKWSKMGRRMRYCWSWARMMYYAGSYMCKRSRWGCGY 51420
QY 862 GTATATCGGGGATACGAAGATATTTCTGCAGCGCTGGAATATTTGCTGAGGAAAAAGAAAT 921
Db 51419 KGRSRSRSGMSWRTRGYSYARCYMSKAGAGCKGA - GSYAKGSAGWRWGSYKKGAWY 51361
QY 922 CCTAGCGCGCCAGACCCCTCGTCAGATAGCGTCCCGCAAGAAATTCCCGCTGTAACCAAG 981
Db 51360 SCRMSASKSRMTCYWSCCCTSCGKCYWYMSWKYGAGMCWYMRWCTCMARAMWRARMA 51301
QY 982 AAAGCGGAAGGGCGCACCCCGGACGAGAAAGCAGGAAAAAGAGGCCCTCCAGAAGAC 1041
      | : : : : | : : : : : : : : : : | | | : | : : : : : : : : : :
Db 51300 AMRRARMAAGARAWMSMWMMSMMMMRBHAAAMARMAAAKCMWYWSYRKYCWMSMMMAS 51241
QY 1042 TCGGAGGACGACATGCAGGACGAGGCTTCTGAGAAAATCTGCGCCGCTCCCGGAAGAC 1101
      : | : : : | : : : : : : : : : : | | | : | : : : : : : : : :
Db 51240 WAAMAMMMMTAAWYKKWTIRMRWSKYSSSYAYAWYATMCWICYSSMYAARRMGTYSGG 51181
QY 1102 GACGAAGTCCCCGAGGACACCGACGA 1127
      : : : : | : | : : : : : : : : : : : : : : : : : :
Db 51180 KRRKWAISTCMCCCRRGCCMKKKGMR 51155

RESULT 12
LOCUS SCYOR053W/c 3022 bp DNA linear PLN 11-AUG-1997
DEFINITION S.cerevisiae chromosome XV reading frame ORF YOR053w.
ACCESSION 274961 Y13140
VERSION 274961.1 GI:1420188
KEYWORDS
SOURCE
  baker's yeast.
  Saccharomyces cerevisiae
  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
  1 (bases 1 to 616)
  Landt, O., Hiesel, R. and Unseld, M.
  JOURNAL
  Unpublished
  2 (bases 1 to 3022)
  Bohn, C., Bolotin-Fukuhara, M., Daignan-Fornier, B., Dang, D. V. and
  Valens, M.
  JOURNAL
  Unpublished
  3 (bases 1 to 3022)
  MIPS.
  Direct Submission
  TITLE

```


gene
complement(4666. .6690)
/gene="YOR29-05"
CDS complement(4666. .6690)
/gene="YOR29-05"
/codon_start=1
/protein_id="CAA94539.1"
/db_xref="GI:2104866"
/translation="MTNKSLSLKNRKGVASNTLSGAEQANIGSSAMPDNTSTGPFSSV
SSLDTPVYVRKSTSPGTSQTSKINNASGTSGAVSNTPPEGLKRIPTVTFSDPKLSLR
SILSNADMGNNISLAKKLSFTGGTDSLNSDSSDSPRKEHPHYVEDPLHTPSVR
SRNSTSPRPSVVNTFNPINIEREGISKTGEPQLLESVLEEAMSPNAVSNPLKREN
IMTNDPRLPDDGKLHYLFGATGSLSVFKLMIRKLEEIYGRDKICIQVILNLSAT
TNKQEEKQMASTGFPSTLGGSRTYSNSNAGNITGNKKVASLEKFSIOKTSNSASQ
RQRTDPLHIELRMADILVAPLTANTILAKIALGLCDNLITSVIRAMNPTPEPLAP
SMGSGTFNSIMTKKHFRIQEEPMWTVFKPSEKVMGINGDIGSGMDANEIVGKIY
VKLGYPDVSAKKEEDEDNDEEDNKNKNTGKDEDNDDDDDDDDDDDDDDDDDDDDD
6311. .6745
/gene="YOR29-06"
6311. .6745
/gene="YOR29-06"
/codon_start=1
/protein_id="CAA94540.1"
/db_xref="GI:2104867"
/translation="MCTDVAFFSLDCIATWLGVCNSTDLRLPNFESLNTVIGILENP
GSGVLLTFAEVPVLAFMIDLVCEPVLVDFTTGVSRLETEENGPFVEFVSGIADEPIL
ACSAPEPVFAATPPLFFRLDLFVIASLHYNYGQYSSFWSL"
complement(7080. .8459)
/gene="YOR29-07"
complement(7080. .8459)
/gene="YOR29-07"
/codon_start=1
/protein_id="CAA94541.1"
/db_xref="GI:2104868"
/translation="MTENQTAHVRAILIDATPLITQSYTHYONYAQSEYTTPTVFOEI
KDAQARKNLEIWQSLGTLKLVHPSENSIAKSTFAKLTGDYSLVANDLHILALTYEL
EIKLNGDWRLRKPGDALDASKADVTDGKQKLTEDNKKKEESESVPKKKNRRGK
KOKAKREAREAREANANLEESKAEHVEAGSKEQICNDENIKESSDLNEVEDAD
DDGDWITPENLFEALIKDSGEDTTSGLVEASEEDRHVALNRPENOVALATGFAVON
VALQMNLMNFMNSGLIKIRIRNYMLRCHACFKIEPLPKDGKPKHFCACGGQGTLLR
CAVSVDSTRGNVTPHLKSNFQWNNRCNRYSVASPLSKNSQKRYGKGHVHSPQENVI
LRDQKEYEKVIKQEEWTRRHNEKILNMWIGGSADNYISPFAITGLKQHNVIIGKGR
YVNSSKRRS"
9019. .10206
/gene="YOR29-08"
9019. .10206
/gene="YOR29-08"
/codon_start=1
/protein_id="CAA94542.1"
/db_xref="GI:2104869"
/translation="MPVEKDLKTAYKALYDEKEPLKALHLYDEILKGSPTNLALIFK
AACLEKLYFGFSDWHSDATMENAKELLDKALMTAEGRDRSKIGLYNERFYVHEFNK
DYELAQSYFKKAKNLGYVDDTLPLMEDRLKLNKKKKQKDSINKHTIKPVEESIENR
GDNSSHSPIPLKIEIAPQESPKFKIDWQSTSVTISLFTVMPESEKQVNIYISP
NDRRTLISYQVPKSGSEFOYNAKLSHEVDPAKAVSLKIFPKLEITLISKIDSTQMKRL
EEDILTESSRLSDEGRKNSDSATRLLSAETASKERLSYPPSSKKKIDMSKTDIDEADE
EAGSADSFQKLYAGADPDTKRAMMKSFTESNGTALSTDWEDVSKGTVKTSPPEGMER
KHW"
complement(10521. .13178)
/gene="YOR29-09"
CDS complement(10521. .13178)
/gene="YOR29-09"
/codon_start=1
/protein_id="CAA94543.1"
/db_xref="GI:2104870"
/db_xref="SWISS-PROT:P50275"

/translation="METATSSPLPIKSRNSSENSGTTVIPHNNPSLATPLTVSTWYN
QNSKEPFMKLTTPVRIIRDEGSLPKNVSTNYHFLDSENGKGNMDNMYRENFILISKDLE
KLEENLNIYQNIIGYSNTEIITKEKIITTTISNSIKOFFEQADEELKRLSAENGIEQD
ILNNILERINDPSGIKTIPDLYIRNALILQESKTVPOSPPKPLSLSKKAALDTAKKE
VLGSELPRLRDYLUKSLITLKLHIOQVKNLPGLEADNEAIAEPELSTLAYLLQIE
NGKDIGLSMKFIIDNRKDIKGSARKTINEESVKHMEVIAKIEYEYERRKSVLTK
KVISISSICEQLGTPLATLIGEDFEQDILRSYGEENSTSEIPNFHPVDRERMSKIDITL
EKLQAIHKERADKRLIMEQCKLWRLKISQEYIKTFMRNNSLSTESIGRISKEVM
LETCENELKRLKEELTYLKPILKLSIDESLQEDQEFLEERSKDSRLSRNSHKILL
TEEMKRRIIRHEPRVINDLRIKLEBADGLFDQPLFKGKPLSEADITQOEIEAKYP
RCRVRMQRSKKGCGAKENKVIKNTFKATESIRIYIGLNLNDANITYKTPSKTTIQ
GLTKNDLSQENSLARHMGGTTRKLSPPNRATRLAPTVISRNSKGNTERPTLNRNRS
DLSSSPRINHTGHEHAVKPROLEPIPLNKVDTKGSHIPOLTEKALELLKRSTGTTGK
ENVRSPERKSLEDYAOKLSSPYKEPEHSIYKLSMSPGKFQNLNQKDIESGFDTS
MMEDENDKDFITWKNQVSKLNGFSFTDI"
15476. .15574
/gene="tRNA-Lys"
/product="tRNA-Lys"
15476. .15574
/gene="tRNA-Lys"
complement(15739. .17091)
/gene="YOR29-10"
complement(15739. .17091)
/gene="YOR29-10"
/codon_start=1
/protein_id="CAA94544.1"
/db_xref="GI:2104871"
/translation="MTSDKHLFVLHGLMGNTYTHMESNRTILSTLKKEDVNDMIYE
LPKQNA MFKTFDGEIIEIGYRTLIECEEIRDYKDGKITKLSVMGYSGGLVARFMIGK
MLTEFKELFEDIEPOLFITMATPHGVEEYFNPTGIAYKSALYSALRTUGSTIIGKSGR
EMFIANSNNLLVKLSQGEYLEALSLEFKWRIAFANYKNDRTVAFTATITDCDFIDE
DNKLKYTEFEKIPGSGYKGLRPIVDNALNVNSAPTKPTKTYKKWGRITLILVAT
FLILPILALVMNGLGTAYSIVYCKYRKMLNSGILNEVRKGLGTEQLKGYVTDAYGS
IINSALMDANDANYEASNSNLVNEELPWKEFIQKTYTTINDGVKSKFRKLPEDENRKVI
LRNLNKLKIRVPIYIKAVNAHGVIVARGMDENTAATGIACIEFTAQLLAYLMHKS
N"
complement(17223. .17996)
/gene="YOR29-11"
complement(17223. .17996)
/gene="YOR29-11"
/codon_start=1
/protein_id="CAA94545.1"
/db_xref="GI:2104872"
/translation="MSRKLCTLNFETLSGKQGLVIRDIQLSMNRPTASKSTSELRGOF
IQYVDLAKLPLWVRSTNMNTYRCYSATQAQAYFKSKLRNANRGIVIELFDKVDORSO
EPAYLIIFRENTFLNCFQVDLTMKHEHFDGQVTKLKQDICKTRASVSKEGSDIITIQQ
"

Query Match 3.3%; Score 43; DB 8; Length 54719;
Best Local Similarity 56.8%; Pred. No. 2.9;
Matches 79; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY	1036	GAAGACTCGAGAGCAGCATGCAGGAGGCTTCTGAGAAAAATCCTGCCGCCCTCCCC	1095
Db	4884	GAAGATAATGACGACGACGACGATGATGATGATGACGATGATGACGACGACGAC	4825
QY	1096	GAAGACGACGAAGTCCCGGAGGACACCGACGATGATCCAACTCGGATCCTGACTAT	1155
Db	4824	GACGACGACGACGACGACGACGATGATGATGATGATGATGATGATGATGATGATGAT	4765
QY	1156	TACAATGACATGCCCGCCG	1174
Db	4764	GATGACGACGACGACGACG	4746

RESULT 14
ARI35147
LOCUS ARI35147 390 bp DNA 1linear PAT 16-MAY-2001
DEFINITION Sequence 7 from patent US 6194550.
ACCESSION ARI35147
VERSION ARI35147.1 GI:14124052

[illegible]

```

RESULT 15
AP004592/c
LOCUS
DEFINITION
AP004592 149333 bp DNA linear HTG 27-DEC-2001
Oryza sativa chromosome 8 clone P0666G10, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
AP004592
AP004592.1 GI:18146739
HTG: HTGS_PHASE2.
Oryza sativa (cultivar:Niponbare) DNA, clone:P0666G10.
SOURCE
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 149333)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (26-DEC-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1..149333

```

FEATURES	Location/Qualifiers
source	1. 149333
	/organism="Oryza sativa"
	/cultivar="Nipponbare"
	/db_xref="taxon:4530"
	/chromosome="8"

```

/clone="P0666G10"
BASE COUNT      42214 a 32790 c 32876 g 41351 t 102 others
ORIGIN
Query Match      3.2%; Score 41.4; DB 2; Length 149333;
Best Local Similarity 49.3%; Pred. No. 8.5;
Matches 108; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 276 CCGTCGCGCGCTCGACGCGCTAACCATACCGGCGTTGGCCCGTATAACAGATACCTCAC 335
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109125 CCTTCGCGCGCTCATGCCGCGCAACGCTCCGGCGCGCGCGCGCGCGCTCCC 109066
QY 336 TAGGGTATCAAGAGCGCTGCGACGTTGTCTGAGCTCAACCCGATTTCTAACGTGAGACAT 395
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109065 AGCGGCGGAGGACGACGACGACGACGACGATGGCGCTCGACGACGTCGCTGCCCTCTCCGAGCT 109006
QY 396 GATATCGCGCGGCCCAAGAAAAGAGAAGGGGGGGCCCTTTCGAGGCCCTCCGTCTGTT 455
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109005 GGGCACCTTGTTCCAAGAAAGTCTGACGACGACGACGACGACGCGCGCGCTACCCCGACTGGAC 108946
QY 456 CTACGTGATTTAAGGGCGACGACGCGCGAGGACAAGTACTG 494
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108945 ACACCTCGATGTTTCGACGACGACGACGACGACGCGCGAGGGCGG 108907

```

Search completed: September 29, 2002, 01:37:24
Job time: 10852 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2002, 00:13:43 ; Search time 238.59 Seconds
(without alignments)
9390.894 Million cell updates/sec

Title: US-09-994-064-10
Perfect score: 1305
Sequence: 1 ATGCACCGTCTCATCTCAG.....TAAATGCGCGGTAGCTAA 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1305	100.0	1305	15	AAQ68944	gd gene of infecti
2	1305	100.0	1305	16	AAT33505	Infectious laryngo
3	1305	100.0	1305	20	AAx81152	Seq ID No: 19 of U
4	1305	100.0	1305	21	AAC67867	Recombinant fowlp
5	1305	100.0	1305	21	AAZ49300	ILTV glycoprotein
6	1305	100.0	13473	16	AAT33504	Infectious laryngo
7	1305	100.0	13473	17	AAT44384	Infectious laryngo
8	1305	100.0	18912	17	AAT44385	Infectious laryngo
9	361	27.7	3502	13	AAQ22986	Sequence of ILTV g

10	42	3.2	42	17	AAT44399	Homology vector 56
11	41.6	3.2	390	13	AAQ21833	Randomising oligon
12	41.6	3.2	390	14	AAQ36859	PCR primer for 5'
13	41.6	3.2	390	22	AAf76910	Sequence containin
14	38.4	2.9	1761	21	AAC44461	Zea mays DNA fragm
15	38.2	2.9	1232	23	ABLI1609	Drosophila melanog
16	38.2	2.9	3038	23	ABLI16776	Drosophila melanog
17	38.2	2.9	3232	19	ABLI1608	Drosophila melanog
18	37.4	2.9	2277	19	AAV13834	Homo sapiens ambig
19	37.4	2.9	2277	19	AAV05370	Human telomerase p
20	36.8	2.8	1450	12	AAQ13227	PRV glycoprotein g
21	36.8	2.8	1450	12	AAQ13408	pseudorabies virus
22	36.8	2.8	3395	21	AAD00601	Human membrane tra
23	36.2	2.8	2943	17	AAT16480	SA8 virus gb glyco
24	35.6	2.7	429	20	AAZ27978	Human lectomedin-1
25	35.6	2.7	429	21	AAA96005	Human lectomedin-2
26	35.6	2.7	810	23	ABLI14835	Human lectomedin-1
27	35.6	2.7	1680	23	AAS68538	DNA encoding novel
28	35.6	2.7	2428	22	AAS29596	Human endocrine po
29	35.6	2.7	3642	21	AAA96040	Human lectomedin-2
30	35.6	2.7	4091	21	AAA96039	Human lectomedin-2
31	35.6	2.7	4422	21	AAZ87685	Human G protein-co
32	35.6	2.7	5610	20	AAZ28008	Human lectomedin-2
33	35.6	2.7	5610	21	AAA96035	Human G protein-co
34	35.6	2.7	5610	21	AAA96041	Resequenced human
35	35.6	2.7	5659	21	AAZ87686	Human G protein-co
36	35.4	2.7	1364	14	AAQ35092	Grass pollen aller
37	35.4	2.7	3489	21	AAA30290	Kaposi's sarcoma-a
38	35.4	2.7	3489	22	AAF82901	Nucleotide sequenc
39	35.4	2.7	32207	20	AAV73805	KSHV LTR DNA (nucl
40	35.4	2.7	137507	19	AAV19941	KSHV long unique c
41	35.2	2.7	150	20	AAV64956	Mouse histone H2B
42	35.2	2.7	625	21	AAC45074	Arabidopsis thalia
43	35.2	2.7	707	24	ABLI01484	Murine apoptosis r
44	35.2	2.7	1508	21	AAC43746	Zea mays DNA fragm
45	34.6	2.7	1194	21	AAZ51529	S. paucimobilis am

ALIGNMENTS

RESULT 1	
AAQ68944	
ID	AAQ68944 standard; DNA; 1305 BP.
XX	
AC	AAQ68944;
XX	
DT	13-APR-1995 (first entry)
XX	
DE	gd gene of infectious laryngotracheitis virus (ILT).
XX	
KW	Homology vector 586-36.6; gd gene; cassette gene; ss.
XX	
OS	Infectious laryngotracheitis virus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	1.1305
PN	/*tag= a
PN	WO9419014-A.
XX	
PD	01-SEP-1994.
XX	
PF	28-FEB-1994; 94WO-US01826.
XX	
PR	26-FEB-1993; 93US-0024156.
XX	
PA	(JAPG) NIPPON ZEON KK.
XX	
PI	(SYTR) SYNTRON CORP.
XX	
PI	Cochran MD;
XX	
DR	WPI; 1994-294007/36.

DR P-PSDB; AAR58856.
XX
PT New recombinant fowl pox virus for use in vaccines - contains
PT genes expressing antigens of Newcastle disease virus and opt.
PT infectious bronchitis virus
XX
PS Disclosure; Page 74-75; 85pp; English.
XX
CC AAQ68944 contains the coding region of the ILT gd gene and was
CC derived from an approx. 2060 bp EcoRI to BclI restriction sub-
CC fragment of the ILT Kpni genomic restriction fragment #8 (10.6kb).
CC It is part of a cassette in plasmid 586-36.6 constructed for
CC the purpose of inserting ILT gb and gd genes into the fowlpox
CC virus (FPV).
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

Query Match 100.0%; Score 1305; DB 15; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGCTCCTCATCTCAGACGGGACCTCCGTTACTACGGCGAAGAGAGAGGTGCTTAAC 60
Db 1 atgcacgctcccatctcagacgycgactcgcgttactacgcaagagagagtgcttaac 60
QY 61 AAACACATGATGTCGGGTGGAAGAACGGTCTGCTCAGGGCGCAGCTGTATTCACTCTTTTC 120
Db 61 aaacacatgattgctgggtggaagaaacggtctgctcagggcgagctgtattcactctttc 120
QY 121 TGGACTTGTGTGACGATTATGCGGGACATATCTGCTTTGTAGCGCAACGCTATGACCGC 180
Db 121 tggacttgtgtgacgattatgcgggacatattctgctttgtagcgcaacgctatgacggc 180
QY 181 CATTTATTTTGGAGGAATGCTTTTGGACTATCGTACTGCTTCTTCCCTTGCTAGCCAG 240
Db 181 catttattttgaggaatgcttttggactatcgtaactgcttcttcccttgctagccag 240
QY 241 AGCACCGCCGCGCTCACGCTACGACTACATTTTAGCGCGCTCGCGCTGACGGCGCTAAC 300
Db 241 agcacccgcccgcgtcacgctacgactacattttagcgccgtcgcgcgtcgacgcgctaacc 300
QY 301 ATACCGGCGGTTGGCCGCTATACAGATACCTCCTACTAGGGTATCAAGAGCGCTGCGACGTT 360
Db 301 ataccggcggttgcccgctataacagataccctcaactaggtatcaagagcgctgcgacgtt 360
QY 361 GTCGAGCTCAACCCGATTTCTTACGTGACGACATGATATCGGGCGCCAAAGAAAAGAG 420
Db 361 gtcgagctcaacccgatttctaactgtgacgacatgatatcgcgcccaagaaaagag 420
QY 421 AAGGGGGGCCCTTTCGAGGGCCCTCCGTCGTGTGTTCTACGTGATTAGGGCGACGACGGC 480
Db 421 aaggggggcccttcgagggccctccgctcgtgtgttctacgtgattaaggcgacgagcg 480
QY 481 GAGGACAAGTACTGTCCCAATCTATAGAAAAGAGTACAGGGAATGTGGCGACGTACAACTG 540
Db 481 gaggacaagtactgtcccaatctatagaaaagagtlacaggaatgtggcgacgtacaaactg 540
QY 541 CTATCTGAATGCGCCGCTCAATCTGACACAGATGTGGGCGAGTGAATGTTCTTCCAGACC 600
Db 541 ctatctgaatgcccgcgtcaatctgacacagatgtgggagtgactatgttccctagacc 600
QY 601 CTGTATCGCGAAATGGCGCGGAGCTGACTATATTCGCCCACTGCTGCGCTCTCTGGC 660
Db 601 ctgtatcgcgaaatggcggcgagctgactatatctcccccactgtgcgtctctgyc 660
QY 661 CAATFACCTGCTGACCCCTGAAAATCGGGAGATTGGCAAAACAGCTCTGTAACCTCTAGAA 720
Db 661 caatfacctgctgacccctgaaaatcgggagattggcaaaacagctctgtaacctctagaa 720
QY 721 GTTAAAGATCGCTGTTTAAAGATCGGGGTGCGACGCTTAACCTTTTACCGTCGAAATGCTGG 780
Db 721 gttaaagatcgctgtttaaagatcggggtgcgacgcttaaccttttacccgtcgaatgtg 780

QY 781 ACAACAGAACAGTATCAGACTGGATTTCAGGGCGGAACACCTTTATCCGATCGACACACC 840
Db 781 acaacagaacagtatcagactggatttcagggcggaacacctttatccgatcgacacacc 840
QY 841 AATACAGACACGCGGACGACGATATTCGGGGATACGAAGATATTCTGACGGCTGGAAT 900
Db 841 aatcacgacacgcyggaacgctatatctcgggatacgaagatatctgcagcgcgtgaa 900
QY 901 AATTTGCTGAGGAAAAGAAATCCTAGCGCCCGCAGACCTCTGTCAGATAGCTCCGCAA 960
Db 901 aattgctgagaaaaaagaatcctagcgcgcagaccctcgtccagatagcgtccgcaa 960
QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGACCGCACCCCGGACGCAAGAAAGCAGCAA 1020
Db 961 gaattccgcgtgtaaccaagaagcggaaagcgccaccccgagcgcgaagacagcga 1020
QY 1021 AAGAAGGCCCTCCAGAAAGACTCGGAGACGACATGACGGCAGAGGCTTCTGAGAAAT 1080
Db 1021 aagaagccctccagaaagactcggagacgacatgcagcagagagcgtctcgggaaat 1080
QY 1081 CCTGCCGCCCTCCCGAAGACGACGAAGTCCCGGAGACACCGAAGCAGATGATCCAAC 1140
Db 1081 cctgcgcgcctcccgaaagacgaagtcgccgagacacccgagacacgacgatccaac 1140
QY 1141 TCGGATCCGACTATTACAAATGACATGCCCGCGGTGATCCCGGTGAGAGAGACTACTAA 1200
Db 1141 tcggatccgactattacaatgacatgcccgcggtgattcccggtgagagagactactaa 1200
QY 1201 AGTTCTAATGCCGCTCTCCATGCCCATATTTCGGCGCGTTCGTAGCCTGCGCGCTC 1260
Db 1201 agttctaattgccgctctccatgcccatatttcggcgcggttcgtagcctgcgcgctc 1260
QY 1261 GTGGGCTACTGTTTGAGACATCGTAAATGCGCGCTAGCTAA 1305
Db 1261 gtgggctactgtttgagacatcgtaaatgcgcgctagctaa 1305

RESULT 2
AAT33505
ID AAT33505 standard; DNA; 1305 BP.
XX
AC AAT33505;
XX
DT 19-NOV-1996 (first entry)
XX
DE Infectious laryngotracheitis virus gd gene.
XX
KW Infectious laryngotracheitis virus; ILTV; herpesvirus;
KW attenuation; vector; vaccine; chicken; poultry; immunisation;
XX glycoprotein gd; ds.
XX
OS Infectious laryngotracheitis virus.
XX
PN WO9508622-A1.
XX
PD 30-MAR-1995.
XX
PF 16-SEP-1994; 94WO-US10628.
XX
PR 24-SEP-1993; 93US-0126597.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Cochran MD, Wild MA;
XX
DR WPI: 1995-139591/18.
DR P-PSDB; AAW00638.
XX
PT Recombinant attenuated infectious laryngotracheitis virus - for use
PT in vaccines to protect poultry from infection from the virus, also
PT methods of distinguishing between vaccinated and naturally infected
PT birds

XX Example 1; Page 106-107; 177pp; English.
PS
XX The gd gene (AAT33505) spans bases 8462-9766 of the unique short
CC region (see also AAT33504) of infectious laryngotracheitis virus
CC (ILT) genomic DNA. It codes for a glycoprotein (AAW00638) of
CC approx. 48,477 mol.wt. that is homologous to pseudorabies
CC virus g50 and to gd from herpes simplex virus-1, Marek's
CC disease virus, IPV and bovine herpesvirus-1.1. Monoclonal
CC antibodies raised to ILTV react specifically with gd from
CC ILTV and also with ILTV gd expressed in herpesvirus of
CC turkeys (HVT) virus vector. ILTV gd expressed in the HVT
CC vector is useful as a subunit vaccine.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

Query Match 100.0%; Score 1305; DB 16; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCGTCTCATCTCAGACGGCACTCGCGTTACTACGCGAAAGAGAGGTGCTTAAC 60
Db 1 atgcaccgctccatctcagcgcgactcgcgttactacgcgaaagagagtgcttaac 60
QY 61 AAACACATGATGCGGTGGAACCGTCTGCTCAGCGCGCAGCTGTATTCACTCTTTTC 120
Db 61 aaacacatgatgtgcgtggaacacggtgctgctcagcgcgagctgtattcactctttc 120
QY 121 TGGACTTGTCAGGATATGCGGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
Db 121 tggacttgctcaggatattgctgggagcatatctgctttgtacgcaacgctatgacccgc 180
QY 181 CATTTATTTTGAGGAATGCTTTTGGACTATCGTACTGCTTTCTCTCGCTAGCCAG 240
Db 181 catttattttgaggaatgcttttggactatcgctactgctttcttccctcgtagccag 240
QY 241 AGCACCGCGCGCTCAGCTACGACTACATTTTGAAGCGCGCGCTCGACGGCTAAC 300
Db 241 agcacccgcgctcagctacgactacatttagcgctgcgcgctcgacgcgctaacc 300
QY 301 ATACCGCGGTTGGCCGTATAACAGATACCTCAGTAGGCTATCAAGAGGCTGCGAGCTT 360
Db 301 ataccgcggttgcccgatatttaacagatgacacatgatacgcgccaaagaaagag 420
QY 361 GTGAGCTCAACCCGATTTCTAACGTTGAGCAGACATGATATCGCGGCCAAAGAAAAGAG 420
Db 361 gtgagctcaacccgatttctaacgttgagcagacatgatacgcgccaaagaaagag 420
QY 421 AAGGGGGCCCTTTCGAGGCCCTCCGTCGTCGTGTTCTACGTGATTAAGGGGAGCAGCGGC 480
Db 421 aagggggccctttcgaggccctccgctcgtctgttctacgtgattaagggcgagcagcgc 480
QY 481 GAGGACAAGTACTGTCCATCTATAGAAAAGAGTACAGGGAATGTGGCGAGCTACAACTG 540
Db 481 gaggacaagtactgtccaatctatagaaaagagtagcaggaatgtggcgactacaactg 540
QY 541 CTATCTGAATGCGCGTTCAATCTGCACAGATGTGGCAGTGAAGTATGTTCTTAGACAC 600
Db 541 ctatctgaatgcgcgcttcaatctgcacagatgtggcagtggaactatgttccctagacc 600
QY 601 CTTGTATCGGGAATGGCGGGGAGCTGACTATATTCGCCCACTGCTGCGCTCTGCGC 660
Db 601 ctgtatcggaatggcggggagctgactatatctccccaactgtcgcctctcggc 660
QY 661 CAATACCTTGCTGACCCGTGAATAATCGGGAGATTGCGCAACAGCTCTCGTAACTGTAGAA 720
Db 661 caatacttgctgacccgtgaataatcgggagatttgcgcaaacagctctcgtaaacttagaa 720
QY 721 GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAATTTTACCCTGGAATGCTGG 780
Db 721 gttaacgatcgcgtgtttaaagatcgggtcgcagcttaacttttacgcgtgaaatgctgg 780

QY 781 ACAACAGACAGTATCAGACTGGATTTCAGAGCGGAACACCTTTATCCGATCGCAGACACC 840
Db 781 acacagacagatcatcagactggtattcaagcggaacaccttaccgatcgcagacacc 840
QY 841 AATACAGCAGACGCGGACGACGTAATTCGGGATACGAAGATAATTCTGACGCGCTGGAAT 900
Db 841 aatacagcacgcgcgacgacgtatatacgggatacgaagatatcttcgacgcgtggaat 900
QY 901 AATTGCTGAGGAAAAAGAAATCCTAGCGCGCCAGACCCCTCGTCCAGATAGCGTCCGCCAA 960
Db 901 aattgctgaggaaaaagaatcctagcgcgcagaccctcgtccagatagcgtccgcgaa 960
QY 961 GAAATTCCTCGCTGTATACCAAGAAAGCGGAAGGGCGCACCCCGGACGCAAGACGCGAA 1020
Db 961 gaatctccgctgttaaccaagaagcggaaaggcgaccccgagcgcgaagcagcgaa 1020
QY 1021 AAGAAGCCCCCTCCAGAAAGACTCGGAGACGACGATGACGAGCAGCGCTTCTGAGAAAT 1080
Db 1021 aagaagccccctccagaagaactcggagacgacatgcagcgagagcttctgagaaat 1080
QY 1081 CCTGCCGCCCTCCCGAAGACGACGAAGTCCCGAGGACACACCAGCAGCATGATCCAAAC 1140
Db 1081 cctgccgccctcccgaaagacgaagatcccgaggaacacggaagcagatgatccaaac 1140
QY 1141 TCGGATCCTGACTATTACAATGACATGCCCGCCGCTGATCCCGGTGAGAGACTACTAA 1200
Db 1141 tcggatcctgactattacaatgacatgcccgccgctgatccgggtgagagactactaaa 1200
QY 1201 AGTTCTAATGCCGTCTCCATGCCCATATTCGGCGGCTTCGTAGCCTGCGGCTGCGGCTC 1260
Db 1201 agttctaattgccgtctccatgcccatattcggcggttcgtagcctgcgcgctgcgctc 1260
QY 1261 GTGGGCTACTGTTTGAGGACATCGTAAATGCGCGCTAGCTAA 1305
Db 1261 gtgggctactgtttgaggacatcgtaaatgcgcgctagctaa 1305

RESULT 3
AAx81152
ID AAx81152 standard; DNA; 1305 BP.
XX
AC AAx81152;
XX
DT 07-SEP-1999 (first entry)
XX
DE Seq ID No: 19 of US5925358.
XX
KW Fowlpox virus; FPV; recombinant; vaccine; immunisation; chicken; NDV;
KW Newcastle disease virus; NDV; Fowlpox; Infectious laryngotracheitis; ds.
XX
OS Fowlpox virus.
XX
PN US5925358-A.
XX
PD 20-JUL-1999.
XX
PF 07-JUN-1995; 95US-0484575.
XX
PR 07-JUN-1995; 95US-0484575.
PR 26-FEB-1993; 93US-0024156.
PR 28-FEB-1994; 94WO-US02252.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Cochran MD, Junker DE;
XX
DR WPI; 1999-418249/35.
DR P-PSDB; AAY21984.
XX
PT Fowlpox viruses, useful as vaccines for immunization of
PT chickens/turkeys against Fowlpox and Newcastle disease virus
XX
PS Disclosure; Columns 77-82; 108bp; English.

XX The invention relates to a recombinant fowlpox virus (FPV) comprising
CC a foreign DNA inserted into a region of the fowlpox virus genome
CC corresponding to a 2.8 kb EcoRI fragment, capable of being expressed in a
CC host cell. The virus is used as a vaccine for immunising chickens against
CC Newcastle disease virus (NDV), Fowlpox, and infectious laryngotracheitis.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

Query Match 100.0%; Score 1305; DB 20; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACGGCAAGAGAGGTGCTTAAC 60
Db 1 atgcaccgtctcatctcagacggcaactcgcttactaagcgaagagaggtgcttaac 60
QY 61 AAACACATGATGGCGTGAAGAAACGGGTGCTCAGGGCAGCTGTATTCACTCTTTTC 120
Db 61 aaacacatgatggcgtgaagaaacgggtgctcagggcagctgtattcaactcttttc 120
QY 121 TGGACTGTGTAGAGATTATGCGGAGCATATCTGCTTTGACGCAACGCTATGACCGC 180
Db 121 tggactgtgtagagattatgcggagcatatctgctttgacgcaacgctatgacccgc 180
QY 121 tggactgtgtcaggtattatgcggagcatalctgttctgacgaacgcatgacccgc 180
QY 181 CATTATTTTGAAGAAATGCTTTTGGACTATCGTACTGCTTTCTCTGCTAGCCAG 240
Db 181 cattattttgaggaatgcttttggactatcgtaactgtaactgttcttcttcctgtagccag 240
QY 241 AGCACCGCGCGCTCAGCTACGACTACATTTTAGCGCGTCCGCGCTCGACGGCTAAC 300
Db 241 agcacccgcgctcagctacgactacattttagcgcgctccgctcgacggcttaacc 300
QY 301 ATACCGCGGTTGGCCCGTATAACAGATACCTCCTAGAGGATCAAGAGGCTGCGAGCTT 360
Db 301 ataccgcggttggcccgatataacagatacctcactagaggtatcaagagctgcgacgtt 360
QY 361 GTGAGCTCAACCCGATTTCTACGTGACGACATGATATCGCGCGCAAGAAAGAGAG 420
Db 361 gtgagctcaacccgatttctacgtgacgacatgatatcgcgcgcaagaaagagag 420
QY 421 AAGGGGGCCCTTTCGAGCGCTCCGCTGCTGTGTTCTACGTGATTAAGGCGACGCGC 480
Db 421 aagggggcccctttcgagcgctccgctgctgtgttctacgtgattaaaggcgacg 480
QY 481 GAGGACAAGTACTGTCCATCTATAGAAAAGAGTACAGGGAATGTGGCGACCTACAAC 540
Db 481 gaggacaagtactgtccatctatagaaaagagtacagggaatgtggcgacctacaactg 540
QY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGCGGACAGTGAATGTTCTTAGCACC 600
Db 541 ctatctgaatgcgcgcttcaatctgcacagatgtgggagtgactatgttctctagacc 600
QY 601 CTGTATCGGGAATGGCGGAGCTGACTATATTTCTCCCGACTGCTGCGCTCTTGCC 660
Db 601 ctgtatcgggaatggcgagctgactatatttctcccgactgctgcgctctctg 660
QY 661 CAATACTTGTGACCTTGAATAATCGGAGATTGCGCAACAGCTCTGTAAGCTTAGAA 720
Db 661 caatacttgtgacctgaataatcgggagatttgcgcaaacagctctgtaactctagaa 720
QY 721 GTTAACGATCGCTTTTAAGATCGGGTTCGACGCTTAACCTTTTACCGTGAAGTCTGG 780
Db 721 gttaacgatcgcttttaagatcgggttcgacgcttaaccttttaaccgtcgaatgtcg 780
QY 781 ACAACAGACAGTATCAGACTGGATTTCAGGCGAACACCTTTATCCGATCGACAGACC 840
Db 781 acaacagacagtatcagactggatttcaggcggaacaccttataccgacgcagacacc 840
QY 841 AATACAGACACGCGGAGACGATATTCGGGATACGAAGATATTCGACGCGCTGAAT 900
Db 841 aatacagacacgcgagacgatatcgggatacgaagataattctgcagcgctggaat 900

QY 901 AATTGCTGAGGAAAAAGAAATCCTAGCGCGCCAGACCCCTGCTCCAGATAGCGTCCGCCAA 960
Db 901 aattgctgaggaanaaagaatcctagcgcgcagacccctgtccagatagcgtccgcgaa 960
QY 961 GAAATTCGCGCTGTAAACCAAGAAAGCGAAGGCGCACCCCGAGCAGCAGAAAGACCGGAA 1020
Db 961 gaatttcgctgttaaccaagaagcgaagggcgacccccgagcagaagcagcgaa 1020
QY 1021 AAGAAAGCCCCCTCCAGAAAGATTCGAGAGACATGCAAGCCAGAGGCTTCTTGAGAAAT 1080
Db 1021 aagaagccccctccagaagatcgaagagacatgcagcagaggttcttgagaaat 1080
QY 1081 CTTGCCGCCCTCCCGAAGACGACGAAGTCCCGAGAGACACCGAGCAGATGATCCAAAC 1140
Db 1081 cttgccgccctcccgaaagacgaagatcccgagagacaccgagcagatgatccaaac 1140
QY 1141 TCGGATCCTGACTATTAACAATGACATGCCCCCGCGGTGATCCCGGTGAGAGACTACTAAA 1200
Db 1141 tcggatcctgactattaaatgacatgcccccgcggtgatcccggtgagagactactaaa 1200
QY 1201 AGTCTTAATGCGGTCTCCATGCCCCATATTCGCGCGCTTCTGATGCGTGGCGGCTGCGCTC 1260
Db 1201 agtcttaatgcggtctccatgccccatattcgcgcgcttctgtagctgcgcgctc 1260
QY 1261 GTGGGGCTACTGTTTGGAGCATCGTAAATGCGCGCGGTAGCTAA 1305
Db 1261 gtggggctactgtttggagcatcgtaaatgcgcgcggtagctaa 1305

RESULT 4
AAC67867
ID AAC67867 standard; DNA; 1305 BP.
XX AAC67867;
AC AAC67867;
XX
DT 02-MAR-2001 (first entry)
XX
DE Recombinant fowlpox virus-related polynucleotide, SEQ ID NO: 19.
XX
KW Fowlpox virus; FPV; antiviral; antibacterial; vaccine;
KW Newcastle's disease; Marek's disease; infectious laryngotracheitis; ds.
XX
OS Unidentified.
XX
PN US6136318-A.
XX
PD 24-OCT-2000.
XX
PF 07-JUN-1995; 95US-0486414.
XX
PR 26-FEB-1993; 93US-0024156.
PR 28-FEB-1994; 94WO-US02252.
XX
PA (JUNK/) JUNKER D E.
PA (COCH/) COCHRAN M D.
XX
PI Cochran MD, Junker DE;
XX
DR WPI; 2000-686071/67.
XX
PT New recombinant fowlpox virus useful as vaccines contains foreign DNA
PT inserted into specific non-essential region of the genome
XX
PS Disclosure; Column 77-80; 56pp; English.
XX
CC The present sequence is provided in a specification relating to a
CC recombinant fowlpox virus (FPV) that comprises a foreign DNA inserted
CC within a 3.5 kb EcoRI fragment of the FPV genomic DNA. The foreign DNA
CC can be expressed in host cells infected with FPV. The recombinant FPV
CC may be used in vaccines to protect animals (especially chickens) against
CC fowlpox and, depending on the source of the foreign DNA, other diseases,
CC particularly Newcastle's disease, Marek's disease or infectious

CC laryngotracheitis.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

Query Match 100.0%; Score 1305; DB 21; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACGCGAAAGAGAGAGTGTCTAAC 60
Db 1 atgcaccgtctcatctcagacgycactcgcgttactacgcgaagagagtgcttaac 60
QY 61 AAACACATGATGCGGTGGAATAACGGTGCTGCTCAGGCGCAGCTGTATTCACTCTTTTC 120
Db 61 aaacacatgattgcggtggaataaacggtgctgcagcgcagctgtattcaactcttttc 120
QY 121 TGGACTTGTCTAGGATTAATGCGGGAGCATATCTGTTGTACGCCAACGCTATGACCGC 180
Db 121 tggacttgtctaggattaatgcgggagcatatctgttgtacgccaacgctatgacccgc 180
QY 181 CATTTATTTTGGAGGAATGCTTTTGGACTATGCTACTGCTTCTCTCTGCTAGCCAG 240
Db 181 cattatttttggaggaatgcttttggactatgctactgcttctctctgctagccag 240
QY 241 AGCACCGCCGCGCTCAGCTACGACTACATTTTAGGCCGCGCGCTCGACGCGCTAAC 300
Db 241 agcacccgcccgtcagctacgactacatctttaggccgcgcgcgtcgacgcgctaacc 300
QY 301 ATACCGGCGGTTGGCCCGTATAACAGATACCTCAGGCTATCAAGAGGCTGCGAGCTT 360
Db 301 ataccgcggttgcccccgtataacagatacctcactagggatatcaagagctgcgacgtt 360
QY 361 GTCGAGCTCAACCCGATTTCTTAACGTGAGCAGACATGATTCGGCGCCCAAGAAAGAG 420
Db 361 gtcgagctcaacccgatttcttaacgtgagcagacatgatatcgcgccaaagaaagag 420
QY 421 AAGGGGGCCCTTTCGAGCGCTCCGTCGTCGTGTTCTACGTGATTAAGGGCAGCAGCGC 480
Db 421 aagggggccctttcgagcgctccgctcgtctgttctacgtgalttaagggcagcagcgc 480
QY 481 GAGGACAAGTACTGTCCAATCTATAGAAAAGAGTACAGGGAATGTGGCAGCTACAACTG 540
Db 481 gaggacaagtactgtccaatctatagaaaagagtacagggaaatgtgagcagctaacgtg 540
QY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGCAGTGGAATGTTCTCTAGCACC 600
Db 541 ctatctgaatgcccgttcaatctgcacagatgtggcagtgagctatgttccctagcacc 600
QY 601 CTTGTATCGGGAATGGCGGGGAGTACTATATTTCTCCCCACTGCTGCGCTCTCTGGC 660
Db 601 cttgtatcggaatggcggggagactgactatatatttctccccactgctgcgtctctggc 660
QY 661 CAATACTTGTGACCCCTGAAATCGGGAGATTGGCGCAACAGCTCTCGTAACCTTAGAA 720
Db 661 caatactgtgacccctgaaaatcgggagatttgcgcaaacagctctcgtlaacttagaa 720
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTGGCAGCTTAATTTTACCCTGGAATGCTGG 780
Db 721 gttaacgatcgctgtttaaagatcgggtgcgagcttaacttttaccgctgaaatgctgg 780
QY 781 ACAACAGAACAGTATCAGACTGATTTCGAAGGGCGAACACCTTTATCCGATCGCAGACC 840
Db 781 acaacagaacagtatcagactgatttcaagggcgaaacactttatccgatcgcacacc 840
QY 841 AATACACGACACGCGGACGAGCTATATCGGGATACGAGATATTTGACAGCGCTGGAAT 900
Db 841 aatcacgacacgcggacgagctatatatcgggatagcgaagatatcttcgacgcgtggaat 900
QY 901 AATTGCTGAGAAAAGAAATCTTAGCGCGCCAGACCCCTCGTCCAGATAGCGTCCGCCAA 960
Db 901 aattgctgagaaaaaagaatccttagcgcgccagaccctcgtccagatagcgtccgcgcaa 960

QY 961 GAAATTCGCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGGACCGAGAAAGACGGAA 1020
Db 961 gaaatcccgctgttaaccaagaagcggaaaggcgcaccccgagcagaaagcagcgaa 1020
QY 1021 AAGAAGCCCCCTCCAGAAGACTCCGAGAGCAGACATGACAGGAGAGGCTTCTGAGAAAAAT 1080
Db 1021 aagaagccccctccagaagactcggagagacatgacagcagaaggtctctgagaaat 1080
QY 1081 CCTGCCGCCCTCCCGAAGACGACGAAGTCCCCGAGGACACCGACGATGATCCAAC 1140
Db 1081 cctgccgccctcccgaaagacgaagagtlcccgaggaacccgagcagcatgatccaac 1140
QY 1141 TCGGATCCTGACTATTACAATGACATGCCCGCGTGATCCGGGTGAGAGACTACTAAA 1200
Db 1141 tcggatcctgactattacaatgacatgcccgcggtgatcccggtgagagactactaaa 1200
QY 1201 AGTTCTAATGCCGCTCTCCATGCCCATATTCCGGCGGCTTCGTAAGCCTGCGCGGTGCTC 1260
Db 1201 agttctaatgccgctctccatgcccatattcgcgcggttcgtagcctgcgcgctc 1260
QY 1261 GTGGGCTACTGTGTTGAGACATCGTAATAATGCCGCGGTGACTAA 1305
Db 1261 gtgggctactgtgttgagacatcgtaataatgcccgcggtgactaa 1305

RESULT 5

AAZ49300
ID AAZ49300 standard; cDNA; 1305 BP.

AC AAZ49300;

DT 14-MAR-2000 (first entry)

DE ILTV glycoprotein D (gd) gene.

XX Fowlpox virus; FPV; recombinant; antigenic protein; expression;
KW infectious laryngotracheitis virus; ILTV; Newcastle disease virus; NDV;
KW Marek's disease virus; cytokine; promoter; homologous recombination;
KW homology vector; multivalent; live vaccine; glycoprotein D; ds.
OS Infectious laryngotracheitis virus.

XX Key Location/Qualifiers
FH 1.1305
FT CDS /*tag= a
FT /product= "Glycoprotein D (gd)"

PN US6001369-A.

PD 14-DEC-1999.

PF 07-JUN-1995; 95US-0477459.

PR 26-FEB-1993; 93US-0024156.

PR 28-FEB-1994; 94WO-US02252.

PA (SYTR) SYNTRO CORP.

PI Junker DE, Cochran MD;

DR WPI; 2000-071638/06.

DR P-PSDB; AAY58184.

PT Recombinant fowlpox virus useful as a vaccine for immunizing fowl
PT against Marek's disease, Newcastle disease, Infectious
PT laryngotracheitis Virus and/or fowlpox -

PS Claim 5; Columns 77-80; 56pp; English.

CC The invention relates to a recombinant fowlpox virus (FPV)
CC comprising a foreign DNA inserted into a 4.2 kb EcoRI fragment
CC of the fowlpox virus genome. The foreign DNA is capable of being
CC expressed in a host cell into which the fowlpox virus has been

CC introduced and encodes an antigenic protein. The antigenic protein
CC which may be expressed includes infectious laryngotracheitis virus
CC (ILTIV) glycoprotein B (gB) or glycoprotein D (gD, AAY58184), Newcastle
CC disease virus (NDV) haemagglutinin (HN, AAY58182) or fusion (F) protein
CC (AAY58183) and Marek's disease virus gB or gD. The foreign DNA may
CC alternatively encode a cytokine such as chicken myelomonocytic growth
CC factor (CMGF) or chicken interferon (cIFN). The foreign DNA in the
CC recombinant FPV is under the control of one or more synthetic pox
CC promoters, enabling control of strength and timing of heterologous
CC gene expression. The synthetic pox virus promoters that may be used are
CC based on promoters of the vaccinia virus and include early promoter 1
CC (EP1), late promoter 1 (LP1), EP2 and LP2 (AAZ49291-249294,
CC respectively). The recombinant FPV is generated via homologous
CC recombination between FPV DNA and a homology vector containing the
CC foreign DNA flanked by FPV sequences. The recombinant fowlpox viruses of
CC the invention are used as multivalent live vaccines for immunising fowl
CC against Marek's disease virus, NDV, ILTV and/or fowlpox virus. The
CC present sequence represents the ILTV glycoprotein D (gD) gene.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

Query Match 100.0%; Score 1305; DB 21; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGCTCTCATCTCAGACGGCACTCGCGTTACTACGCCGAAAGAGAGAGGTGCTTAAC 60
Db 1 atgcaccgctccatcatctcagacgycactcgcgttactacgcgaagaagagaggtgcttaac 60
QY 61 AAACACATGATTCGGGTGGAACACGGTGTGCTCAGCGGCAAGCTGTATTCACTCTTTTC 120
Db 61 aaacacatgattgcgggtggaacacgggtgtgctcagcgcaagctgtattcaactcttttc 120
QY 121 TGGACTGTGTGAGGATTAATGCGGAGCATATATGCTTTGTACGCAACGGTATGACCGC 180
Db 121 tggactgtgtgaggatataatgcggagcatatatgctttgtacgcaacggatgacccgc 180
QY 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTCTCTTCCTTCGCTAGCCAG 240
Db 181 cattatttttgaggaatgcttttggactatcgtaactgcttctccttcgctagccag 240
QY 241 AGCACCGCGCGCGTACGATACGACTACATTTTAGGCCGCTCGCGCTGACGCGCTAAC 300
Db 241 agcacccgcccgcgtacgatacgaactacatttaagccgctcgcgcgtcgaacgcgtaacc 300
QY 301 ATACCGGCGGTTGGCCGTATACAGATACCTCACTAGGGTATCAAGAGGCTGCGACGTT 360
Db 301 ataccggcggttgccccgtataacagatacctcaactaggtatcaagagcgctgcgacgtt 360
QY 361 GTCGAGCTCAACCCGATTTCTAACGTGAGCAGCATGATATCGCGGCCCAAAAGAAAAGAG 420
Db 361 gtcgagctcaacccgatttctaacgtgagcagacatgatatcgcgcccaaaaaaagag 420
QY 421 AAGGGGGCCCTTTTCGAGGCCCTCGTCTGTTCTACGTGATTAAGGCGAGCAGCGGC 480
Db 421 aagggggcccttttcgagggccctcgtctggttctacgtgattaaagggcgacgagcgc 480
QY 481 GAGGACAAGTACTGTCCAAATCTATAGAAAAGAGTACAGGGAATGTGGCAGCTACACTG 540
Db 481 gaggacaagtactgtccaatctatagaaaagagtacagggaatgtggcgacgtacaactg 540
QY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGCAGTGAGCTATGTTCCCTAGCACC 600
Db 541 ctatctgaatgcgcccgttcaatctgcacagatgtggcgagtgagctatgttccctagcacc 600
QY 601 CTGTATTCGCGAAATGGCGGGGACTGACTATATTCCTCCCACTGCTGGCGTCTCTGGC 660
Db 601 ctgtatcgcgaaatggcggggactgactatatctcctcccaactgctgcgtctctggc 660
QY 661 CAATACTTGTGCTGACCCCTGAATAATCGGGAGATTTCGCGCAACAGCTCTGTAAGTCTAGAA 720
Db 661 caatacttgtgacctgaaataatcgggagatttgcgcaaacagctctgtaactctagaa 720

QY 721 GTTAAAGATCGCTGTTTAAAGATCGGGTCCGACCTTAACCTTTTACCGTCGAAATGCTGG 780
Db 721 gttaaagatcgctgtttaagatcgggtccgaccttaaccttttaccgtcgaatatgctgg 780
QY 781 ACAACAGAACAGTATCAGACTGGATTTCAGGCGGAACACCTTTATCCGATCGCAGACACC 840
Db 781 acaacagaacagtatcagactggatttcaaggcggaacaccttaccgatcgacagacc 840
QY 841 AATACACGACACGCGGACGACGATATATCGGGGATACGAAGATATTCTGCAGCGCTGGAAT 900
Db 841 aatacacgacacgcggacgacgatatatcgggatacgaagatatctgcagcgctggaat 900
QY 901 AATTGCTGAGGAAAAAGAAATCTTAGCGCGCCAGACCCCTGTCAGATAGCGTCCGCA 960
Db 901 aattgctgagaaaaaagaatcttagcgcgcgacaccctgtccagatagcgtccgcaa 960
QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAAAGGCGCACACCCCGGACGCAAGAACGCA 1020
Db 961 gaaattcccgctgtaaccaagaagcggaagggcgcaaccccgagcgcaagaagcagcgaa 1020
QY 1021 AAGAAGGCCCTCCAGAGACTCGGAGGACGACATGACGACGACAGGCTTCTGAGAAAAAT 1080
Db 1021 aagaagggccctccagaagactcggaggaacgacatgcaggaagaggtctctgagaaaaat 1080
QY 1081 CCTGCCGCCCTCCCGGAGACGACGAAGTCCCGGAGGACACCGACGATGATCCAAAC 1140
Db 1081 cctgccgccctcccggaagacgacgaagtlcccgaggacaccgacagatgatccaaac 1140
QY 1141 TCGGATCTGACTATTACAATGACATGCCCGCGGTGATCCCGGTGAGGAGACTACTAA 1200
Db 1141 tcggatctgactattacaatgacatgcccgcggtgatcccggtgagagactactaa 1200
QY 1201 AGTTCTAATGCCGCTCTCCATGCCATATTTCGCGGGCGTTGTAAGCCTGCGCGTCCGCTC 1260
Db 1201 agttctaatagccgctctccatgccatatcgcgcggttcgtagcctgcgcgctcgcgctc 1260
QY 1261 GTGGGCGCTACTGTTTGGACATCGTAAATGCGCGCGTAGCTAA 1305
Db 1261 gtgggcgctactgtttggacatcgtaaatgccccgctagctaa 1305

RESULT 6
AAT33504
ID AAT33504 standard; DNA; 13473 BP.
XX
AC AAT33504;
XX
DT 18-NOV-1996 (first entry)
XX
DE Infectious laryngotracheitis virus short region sequence.
XX
KW Infectious laryngotracheitis virus; ILTV; herpesvirus;
KW attenuation; vector; vaccine; chicken; poultry; immunisation; ds.
OS Infectious laryngotracheitis virus.
XX
FH Key Location/Qualifiers
FT repeat_region 1..273 /tag= a
FT /function= internal repeat region
FT misc_feature 274..13371 /tag= b
FT /function= unique short sequence of ILTV
FT CDS complement (281..970)
FT /tag= c
FT /label= US2_gene
FT CDS 1059..2489 /tag= d
FT /label= Protein-kinase_gene
FT CDS 2575..4107 /tag= e
FT /label= U147-like_gene

FT CDS 4113..4445
FT /*tag= f
FT /label= ORF4
FT complement (4139..4519)
FT /*tag= g
FT /label= ORF4_reverse_complement
FT misc_difference 4535
FT /*tag= h
FT /note= "base 4535 is given as 's' in the
FT specification"
FT CDS 4609..5487
FT /*tag= i
FT /label= gG_gene
FT CDS 5697..8654
FT /*tag= j
FT /label= g60_gene
FT CDS complement (6948..7826)
FT /*tag= k
FT /label= ORF6_reverse_complement
FT CDS 8462..9766
FT /*tag= l
FT /label= gD_gene
FT CDS 9874..10962
FT /*tag= m
FT /label= gI_gene
FT CDS complement (10617..11150)
FT /*tag= n
FT /label= ORF8_reverse_complement
FT CDS 11159..12658
FT /*tag= o
FT /label= gE_gene
FT CDS 12665..13447
FT /*tag= p
FT /label= ORF10
FT misc_difference 13002..13003
FT /*tag= q
FT /note= "bases 13002-13303 are given as 'ss' in
FT the specification"
FT repeat_region 13372..13473
FT /*tag= r
FT /function= terminal repeat region
XX PN WO9508622-A1.
XX PD 30-MAR-1995.
XX PF 16-SEP-1994; 94WO-US10628.
XX PR 24-SEP-1993; 93US-0126597.
XX PA (SYTR) SYNTRO CORP.
XX PI Cochran MD, Wild MA;
XX WPI: 1995-139591/18.
DR P-PSDB; AAW00630, AAW00631, AAW00632, AAW00633, AAW00634, AAW00635,
DR AAW00636, AAW00637, AAW00638, AAW00639, AAW00640, AAW00641 W00642.
XX
PT Recombinant attenuated infectious laryngotracheitis virus - for use
PT in vaccines to protect poultry from infection from the virus, also
PT methods of distinguishing between vaccinated and naturally infected
PT birds
XX
PS Example 1; Page 79-94; 177pp; English.
XX
CC The unique short region (AAT33504) of infectious laryngotracheitis
CC virus (ILTV) genomic DNA contains genes (see also AAT33505 and
CC AAT33510-13) that are associated with ILTV virulence. A deletion in
CC those genes, esp. the glycoprotein gG gene, glycoprotein gI gene,
CC thymidine kinase gene, US2 gene, UL47-like gene or the glycoprotein
CC g60 gene, will attenuate the ILTV. A gene for a foreign antigen may
CC be inserted into the US2, UL47-like, ORF4, gG, g60 or gI gene to
CC produce a recombinant ILTV. Recombinant ILTV may be used as a

CC multivalent vaccine, esp. for use in poultry.
XX
SQ Sequence 13473 BP; 3390 A; 3582 C; 3547 G; 2951 T; 3 other;
Query Match 100.0%; Score 1305; DB 16; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCACCGTCCTCATCTCAGACGGGCACTCGCGTTACTACGGGAAGAGAGGTGCTTAAC 60
Db 8462 atgcacgcgtccatcatctcagaacgycactcgcgtactactacgcgaagagagtgcttaac 8521
QY 61 AAACACATGATTGCCGTGGAAAAACGGTGTCTCTCAGGGCGCAGCTGTATTCTTTTC 120
Db 8522 aaacacatgattgcgtggaaaaacggtgtctctcagggcgagctgtattcactcttctc 8581
QY 121 TGGACCTGTGTACAGGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
Db 8582 tggacctgtgtacaggattatgCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 8641
QY 181 CATTTATTTTGTAGGAATGCTTTTGGACTATCTGCTTCTTCTTCTGCTAGCCAG 240
Db 8642 cattatttttgaggaatgcttttggactatctgactgtcttctctctcgttagccag 8701
QY 241 AGCACCGCGCGCGTCACTAGACTACATTTTAAAGCGGCGCGCTCGACGCGCTAAC 300
Db 8702 agcacCGCGCGCGTCACTAGACTACATTTTAAAGCGGCGCGCTCGACGCGCTAAC 8761
QY 301 ATACCGGCGGTTGGCCCGTATACAGAGATACCTACTAGGGTATCAAGAGGCTGCCAGCT 360
Db 8762 ataccgCGGTTGGCCCGTATACAGAGATACCTACTAGGGTATCAAGAGGCTGCCAGCT 8821
QY 361 GTCGAGCTCAACCCGATTTCTTAACGTGGACGACATGATATCGGGGCCCAAGAAAAAGAG 420
Db 8822 gtcgagctcaacCGATTTCTTAACGTGGACGACATGATATCGGGGCCCAAGAAAAAGAG 8881
QY 421 AAGGGGGCCCTTTTCAGGCGCTCCGCTGCTGTGCTTACGTTAAGGGCGACGACGGC 480
Db 8882 aaggggCGCCCTTTTCAGGCGCTCCGCTGCTGTGCTTACGTTAAGGGCGACGACGGC 8941
QY 481 GAGGACAAGTACTGTCCAATCTATAGAAAAAGATACAGGAATGTGGCGACGTACAAC 540
Db 8942 gaggacaagtactgtccaatctatagaaaaagatacaggaatgtggcgacgtacaactg 9001
QY 541 CTATCTGAATGCGCGCTTCATCTGCACAGATGTGGGCAGTGAGACTATGTTCTAGCACC 600
Db 9002 ctatctgaatgCGCGCTTCATCTGCACAGATGTGGGCAGTGAGACTATGTTCTAGCACC 9061
QY 601 CTTGTATCGCGAATGCGCGCGGAGCTGACTATATTTCTCCCCACTGCTGCGCTCTTGGC 660
Db 9062 ctgtatCGCGAATGCGCGCGGAGCTGACTATATTTCTCCCCACTGCTGCGCTCTTGGC 9121
QY 661 CAATACTTGTGACCCGTGAAAAATCGGGAGATTGCGCAAAACAGCTCTCGTAACCTAGAA 720
Db 9122 caatacttGTGACCCGTGAAAAATCGGGAGATTGCGCAAAACAGCTCTCGTAACCTAGAA 9181
QY 721 GTTAAGGATCGCTGTTTAAAGATCGGGTGCAGCTTAACCTTTTACCGTGAATGCTGG 780
Db 9182 gttAAGGATCGCTGTTTAAAGATCGGGTGCAGCTTAACCTTTTACCGTGAATGCTGG 9241
QY 781 ACAACAGAAACAGTATCAGACTGATTTCGAAGCGGAACACCTTTATCGGATCGACACACC 840
Db 9242 acaacagaaacagTATCAGACTGATTTCGAAGCGGAACACCTTTATCGGATCGACACACC 9301
QY 841 AATACAGACACGCGGACGACGATATTCGGGATACGAAGATATTCGACGCGCTGGAAT 900
Db 9302 aatacagacacgCGGACGACGATATTCGGGATACGAAGATATTCGACGCGCTGGAAT 9361
QY 901 AATTGTGTAGGAAAAAAGATCTTAGCGCGCGACACCTCTGCTCCAGATAGCGTCCGCAA 960
Db 9362 aattgtGTAGGAAAAAAGATCTTAGCGCGCGACACCTCTGCTCCAGATAGCGTCCGCAA 9421

```
OY 961 GAAATTCCTGCTGTAAACCAAGAAAGCGGAAAGGCGCACCCCGAGCCAGAAAGCAGCGAA 1020
Db 9422 gaaattccgcgtgaaccaagaagcgaagcgccaccccgacgcagaaagcagcgaa 9481
OY 1021 AAGAAGGCCCCCTCCAGAGACTCGGAGGACGACATGACGAGCAGAGGCTTCTTGAGAAAT 1080
Db 9482 aagaagggccctccagaagaactcggagcgacatgcagcgagaggtctctggagaaat 9541
OY 1081 CCTGCCGCCCTCCCGAAGACGACGAAGTCCCCGAGACACCGACACGATGATCCAAC 1140
Db 9542 cctgccgcctcccccgaagacgagagtcgccgagacacgagcagcatgatccaac 9601
OY 1141 TCGGATCCTGACTATTACAATGACATGCCCGCGGTGATCCCGGTGAGAGACTACTAAA 1200
Db 9602 tcgatacctgactattacaatgacatgccgcgtgatcccggtgagagactactaaa 9661
OY 1201 AGTTCATAATGCCGCTCTCCATGCCCATATTGCGCGGCTTCGTAGCCTGCCGCGTCCGCTC 1260
Db 9662 agtctaatagcgctcctccatgcctatctcgcgcgcttcgtagcctgcgcggtcgcgctc 9721
OY 1261 GTGGGCTACTGTTTGAGCATCGTAAATGCCGCGCTAGCTAA 1305
Db 9722 gtgggctactggtttgagcatcgtaaatgccgctagctaa 9766

RESULT 7
AAT44384
ID AAT44384 standard; DNA; 13473 BP.
XX
AC AAT44384;
XX
DT 02-JUN-1997 (first entry)
XX
DE Infectious laryngotracheitis virus unique short region.
XX
KW ILTV; vaccine; vector; attenuation; poultry;
KW avian infectious bronchitis virus; Newcastle disease virus;
KW infectious bursal disease virus of chickens;
KW Marek's disease virus; herpesvirus; ss.
XX
OS Infectious laryngotracheitis virus USDA strain 8302.
XX
FH Key Location/Qualifiers
FT repeat_region 1..273
FT misc_RNA 274..13371
FT polyA_signal 161..166
FT CDS 107..1010
FT TATA_signal 1040..1043
FT TATA_signal 1042..1045
FT CDS 1059..2489
FT TATA_signal 2491..2496
FT polyA_signal 2538..2543
FT polyA_signal 2538..2543
```

```
FT /note= "polyA signal for ORF2"
FT 2575..4107
FT /tag= k
FT /label= ORF3
FT /note= "unique long 47 (UL47)-like gene"
FT 4523..4526
FT /tag= l
FT /note= "TATA signal for ORF4"
FT 4113..4445
FT /tag= m
FT /label= ORF4
FT complement (4139..4519)
FT /tag= n
FT /label= ORF4(RC)
FT 4523..4526
FT /tag= o
FT /note= "TATA signal for ORF5"
FT 4609..5487
FT /tag= p
FT /label= ORF5
FT /note= "glycoprotein gC gene"
FT 4609..4686
FT /tag= q
FT 4687..5484
FT /tag= r
FT 5564..5569
FT /tag= s
FT /note= "polyA site for ORF5"
FT 5653..5658
FT /tag= t
FT /note= "polyA signal for ORF5"
FT 5697..8654
FT /tag= u
FT /label= ORF6
FT /note= "glycoprotein g60 gene"
FT 6987..7727
FT /tag= v
FT /note= "repeat region consists of approx. 23
FT repeats of 30-36 bp"
FT complement (6948..7826)
FT /tag= w
FT /label= ORF6(RC)
FT 8455..8458
FT /tag= x
FT /note= "potential TATA signal for ORF7"
FT 8461..9766
FT /tag= y
FT /label= ORF7
FT /note= "glycoprotein gD gene, alternative start
FT codon at 8633..8635"
FT 9819..9822
FT /tag= z
FT /note= "TATA signal for ORF8"
FT 9874..10962
FT /tag= aa
FT /label= ORF8
FT /note= "glycoprotein gI gene"
FT 9874..9939
FT /tag= ab
FT 9940..10959
FT /tag= ac
FT complement (10617..11150)
FT /tag= ad
FT /label= ORF8(RC)
FT 11069..11072
FT /tag= ae
FT /note= "TATA signal for ORF9"
FT 11159..12658
FT /tag= af
FT /label= ORF9
FT /note= "glycoprotein gE gene"
FT 11159..11212
FT /tag= ag
FT sig_peptide
```


FT mat_peptide 11213..12655
FT /*tag= ah
FT TATA_signal 12483..12486
FT /*tag= a1
FT /*note= "TATA signal for ORF10"
FT CDS 12665..13447
FT /*tag= aj
FT /*label= ORF10
FT repeat_region 13372..13743
FT /*tag= ak
XX
PN WO9629396-A1.
XX
PD 26-SEP-1996.
XX
PF 21-MAR-1996; 96WO-US03916.
XX
PR 06-JUN-1995; 95US-0468190.
PR 23-MAR-1995; 95US-0410121.
XX
PA (SYTR) SYNPRO CORP.
XX
PI Cochran MD, Wild MA;
XX
XX WPI; 1996-443172/44.
DR P-PSDB; AAW06782;
DR P-PSDB; AAW06783;
DR P-PSDB; AAW06784;
DR P-PSDB; AAW06785;
DR P-PSDB; AAW06786;
DR P-PSDB; AAW06787;
DR P-PSDB; AAW06788;
DR P-PSDB; AAW06789;
DR P-PSDB; AAW06790;
DR P-PSDB; AAW06791;
DR P-PSDB; AAW06792;
DR P-PSDB; AAW06793.
XX
PT Recombinant infectious laryngotracheitis virus with deletion in the
PT glycoprotein G, gI or US2 gene, etc. - useful for vaccines against
PT infectious laryngotracheitis in poultry
XX
PS Example 11; Page 88-103; 216pp; English.
XX
CC The nucleotide sequence of 13,473 bp of contiguous DNA (AAT44384)
CC from the unique short region of infectious laryngotracheitis virus
CC (ILTIV) contains the entire 13,098 bp unique short region and
CC includes 13 open reading frames that encode proteins (AAW06782-94) of
CC over 100 amino acids, 8 of which show significant homology to other
CC virus genes. Novel recomblant, attenuated ILTV comprises the ILTV
CC genome contg. a deletion in the unique short region, esp. in the
CC glycoprotein gG, gI, US2, ORF4, UL47-like or g60 gene. The
CC attenuated virus is useful as a vaccine against ILTV. A foreign
CC gene encoding an antigen e.g. from another avian virus can be
CC inserted into the US2, UL47-like, ORF4, gG, g60 or gI gene to
CC provide a multivalent vaccine for chickens and other poultry.
CC Deletion of the gG or gI gene provides as a negative marker to
CC distinguish vaccinated from infected animals.
XX
SQ Sequence 13473 BP; 3390 A; 3581 C; 3547 G; 2952 T; 3 other;

Query Match 100.0%; Score 1305; DB 17; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACGCACTGCGGTACTACTACGCAAGAGAGAGTGTAAAC 60
Db 8462 atgcacgcgtctcatctcagacgcgcactgcgttactactacgcaagagagagtgcttaac 8521

QY 61 AAACACATGATTCGCGTGGAAACGGTGTCTCAGCGCAGCTGTATTCACTCTTTTC 120
Db 8522 aaacacatgatctgcggtggaaacgggtgtctcagcgcaactgtattcaactctttc 8581

QY 121 TGGACTTGCTCAGGATTATGCGGACATATCTGCTTTGTACCAACGCTATGACCGC 180
Db 8582 tggacttgcctcaggatattgcggaacatactgctttgtacgaacgctatgacgcgc 8641

QY 181 CATTTATTTTGAAGAAATGCTTTTGGACTATCGTACTGCTTCTCTCGCTAGCCAG 240
Db 8642 catttatcttgaagaaatgcttttggactatcgtaactgcttctctcctcgtacgcag 8701

QY 241 AGCACCGCCCGCTCACGTACGATACATTTTAGCCGTGCGCGCTCGACGCGCTAAC 300
Db 8702 agcacccgcccgtcacgtacgatacatttagccgtgcgcgctcgcagcgctaac 8761

QY 301 ATACCGGCGGTGGCCCCGTATACAGATACCTCCTACTAGGATATCAGAGCGCTG 360
Db 8762 ataccggcggttgccccgtataacagatacctcactagggatataagagcgctgcagcgtc 8821

QY 361 GTCGAGCTCAACCCGATTCTTAACGTGGACGACATGATATCGCGGCCCAAGAAAAAG 420
Db 8822 gtcgagctcaacccgattcttaacgtygaacatgatalcgcgcgccaagaaaagag 8881

QY 421 AAGGGGGCCCTTTCGAGGCCCTCCGCTGCTGTTCTACGTGATTAAGGGCGAGCAG 480
Db 8882 aagggggcccctttcgaggccctccgctgctgtgtctacgtgataaaggcgagcagcgc 8941

QY 481 GAGGACAACTACTGTCCAAATCTATAGAAAGAGTACAGGGAATGTGGCAGCTACA 540
Db 8942 gaggacaactactgtccaaatctataagaaagagtacaggaatctggcgactacactg 9001

QY 541 CTATCTGAATGCGCCGTTCATCTGACAGATGTGGCAGTGAATGTCTCTAGCACC 600
Db 9002 ctatctgaatgcccgttcactatctgcacagatgtggcagtgactatgttccctagcacc 9061

QY 601 CTTGTATCCGGAATGGCGCGGAGCTATATTTCTCCCCCCTGCTGCGCTCTG 660
Db 9062 ctgtgatccggaatggcgcggagctatatcttcccccaactgctgcgtctctgctgc 9121

QY 661 CAATACTTCTGACCCCTGAAAAATCGGAGATTTGCGCAACACGCTCTGTAACCT 720
Db 9122 caatacttctgacccctgaaaaatcggagatttgcgcaaacgctctgtaactctagaa 9181

QY 721 GTTAACGATCGCTGTTTAAAGATCGGCTGCGACGTTTAACTTTTACCGTGAAT 780
Db 9182 gttaacgatcgcctgtttaaagatcggctgcgacgttttaaacttttaccgctgaatgctg 9241

QY 781 ACAACAGACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGCAGAC 840
Db 9242 acaacagacagtatcagactgatttcaaggcgacaaccttataccgatacgcagacacc 9301

QY 841 AATACAGCACACGCGGACGACGCTATATCGGGGATACGAAGATATTCTGCAGCGCT 900
Db 9302 aatacagcacacgcggaacgacgactatatacgggatacgaagatacttcgacgcgtgaa 9361

QY 901 AATTGTGTGAGGAAAAAGAAATCTTAGCGCGCCAGACCCCTCTCCAGATAGCGTCC 960
Db 9362 aattgtgtgagaaaaagaatccttagcgcgcagaccctctctcagatagcgtccgcgaa 9421

QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGAAAGGCGCACCCCGCAGCAGCAAGAC 1020
Db 9422 gaatttcgccgctgttaaccaaagaagcgaagggcgcaccccgagcagaaagcagcga 9481

QY 1021 AAGAAGCCCCCTCCAGAGACTCGGAGAGACGACATGCAGGCAAGGCTTCTGAGAA 1080
Db 9482 aagaagccccctccagagactcggagagacatgcagagcgaagggcttctgagaaat 9541

QY 1081 CCTGCCGCCCTCCCGAAGACGACGAAGTCCCGAGGACACCGACAGATGATCCAAC 1140
Db 9542 cctgccgccctcccgaaagacgaagagtcgccgaagacccgaagcagatgatccaaac 9601

QY 1141 TCGATCTGACTATTACATGACATGCCCCGCGTATCCCGTGAAGAGACTACTAA 1200
Db 9602 tcgatctgactatttacaatgacatgccccgcgtgatcccggtggaagagactactaaa 9661

OY 1201 AGTTCATATGCCGCTCTCCATGCCCATAATTCGGCGGCTTCGTAGCCCTGCGCGGTGCGGCTC 1260
|||||
Db 9662 agtctaatgcgcgtctcatgcataatcgcgcgttcgtagcctgcgcgcgttcgcgcctc 9721
|||||
OY 1261 GTGGGCTACTGTTGTGAGCATCGTAATAATGCGCGGTAGCTAA 1305
|||||
Db 9722 gtgggctactggttggagcgcgttaaatgcgcgcgtagctaa 9766
|||||
RESULT 8
AAT44385
ID AAT44385 standard; DNA; 18912 BP.
XX
AC AAT44385;
XX
DT 02-JUN-1997 (first entry)
XX
DE Infectious laryngotracheitis virus unique short + flanking region.
XX
KW ILTV; vaccine; vector; attenuation; poultry;
KW avian infectious bronchitis virus; Newcastle disease virus;
KW infectious bursal disease virus of chickens;
KW Marek's disease virus; herpesvirus; ss.
XX
OS Infectious laryngotracheitis virus USDA strain 8302.
XX
FH Key Location/Qualifiers
FT repeat_region 1..2909
FT /*tag= a
FT /note= "short repeat region"
FT 697..1533
FT /*tag= b
FT /label= SKORF2
FT /note= "US10 gene"
FT complement (2916..3605)
FT /*tag= c
FT /label= SKORF1
FT /note= "short repeat open reading frame 1"
FT 2910..16003
FT /*tag= d
FT /note= "unique short region"
FT 2796..2891
FT /*tag= e
FT /note= "polyA signal for ORF1"
FT complement (2916..3605)
FT /*tag= f
FT /label= ORF1(RC)
FT /note= "unique-like short 2 (US2) gene"
FT 3642..3645
FT /*tag= g
FT /note= "TATA signal for ORF1 and ORF2"
FT 3675..3678
FT /*tag= h
FT /note= "TATA signal for ORF1"
FT 3677..3680
FT /*tag= i
FT /note= "TATA signal for ORF2"
FT 3694..5124
FT /*tag= j
FT /label= ORF2
FT /note= "protein kinase gene"
FT 5126..5131
FT /*tag= k
FT /note= "polyA signal for ORF2"
FT 5173..5178
FT /*tag= l
FT /note= "polyA signal for ORF2"
FT 5210..7081
FT /*tag= m
FT /label= ORF3
FT /note= "unique long 47 (UL47)-like gene"
FT 7158..7161
FT /*tag= n

FT /*note= "TATA signal for ORF4"
FT 6748..7080
FT /*tag= o
FT /label= ORF4
FT complement (6774..7154)
FT /*tag= p
FT /label= ORF4(RC)
FT 7158..7161
FT /*tag= q
FT /note= "TATA signal for ORF5"
FT 7245..8123
FT /*tag= r
FT /label= ORF5
FT /note= "glycoprotein gg gene"
FT 7245..7322
FT /*tag= s
FT 7323..8120
FT /*tag= t
FT 8200..8205
FT /*tag= u
FT /note= "polyA site for ORF5"
FT 8289..8294
FT /*tag= v
FT /note= "polyA signal for ORF5"
FT 8333..11290
FT /*tag= w
FT /label= ORF6
FT /note= "glycoprotein g60 gene"
FT 9623..10363
FT /*tag= x
FT /note= "repeat region consists of approx. 23
FT repeats of 30-36 bp"
FT complement (9584..10462)
FT /*tag= y
FT /label= ORF6(RC)
FT 11091..11094
FT /*tag= z
FT /note= "potential TATA signal for ORF7"
FT 11098..12402
FT /*tag= aa
FT /label= ORF7
FT 12455..12448
FT /*tag= ab
FT /note= "TATA signal for ORF8"
FT 12510..13598
FT /*tag= ac
FT /label= ORF8
FT /note= "glycoprotein gi gene"
FT 12510..12575
FT /*tag= ad
FT 12576..13595
FT /*tag= ae
FT complement (13253..13786)
FT /*tag= af
FT /label= ORF8(RC)
FT 13705..13708
FT /*tag= ag
FT /note= "TATA signal for ORF9"
FT 13792..15291
FT /*tag= ah
FT /label= ORF9
FT /note= "glycoprotein ge gene"
FT 13792..13845
FT /*tag= ai
FT 13846..15288
FT /*tag= aj
FT 15116..15119
FT /*tag= ak
FT /note= "TATA signal for ORF10"
FT 15298..16080
FT /*tag= al
FT /label= ORF10
FT 16004..18912

FT CDS
FT
FT CDS
FT
FT TATA_signal
FT
FT CDS
FT
FT sig_peptide
FT
FT mat_peptide
FT
FT polyA_signal
FT
FT polyA_signal
FT
FT CDS
FT
FT repeat_region

```
FT      /*tag- am
FT      CDS      16129..17013
FT      /*tag= an
FT      CDS      complement (17380..18216)
FT      /*tag= ao
XX      PN      WO9629396-A1.
XX      PD      26-SEP-1996.
XX      PF      21-MAR-1996; 96WO-US03916.
XX      PR      06-JUN-1995; 95US-0468190.
XX      PR      23-MAR-1995; 95US-0410121.
XX      PA      (SYTR ) SYNPRO CORP.
XX      PI      Cochran MD, Wild MA;
XX      PI      WPI; 1996-443172/44.
XX      DR      P-PSDB: AAW06782, AAW06783, AAW06784, AAW06785, AAW06786, AAW06787,
DR      AAW06788, AAW06789, AAW06790, AAW06791, AAW06792, AAW06793, AAW01415,
DR      AAW01416.
XX      PT      Recombinant infectious laryngotracheitis virus with deletion in the
PT      glycoprotein G, gI or US2 gene, etc. - useful for vaccines against
PT      infectious laryngotracheitis in poultry
XX      PS      Example 11; Page 138-154; 216pp; English.
XX      CC      The nucleotide sequence of 19,912 bp of contiguous DNA (AAAT44385)
CC      from the unique short and flanking region of infectious
CC      laryngotracheitis virus (ILTV) contains the entire 13,098 bp unique
CC      short region (see also AAT44384) and includes 17 open reading frames
CC      that encode proteins (AAW06782-94, AAW01415-16) of over 100 amino
CC      acids, 10 of which show homology to other virus genes. Novel
CC      recombinant, attenuated ILTV comprises the ILTV genome contg. a
CC      deletion in the unique short region, esp. in the glycoprotein gG,
CC      gI, US2, ORF4, UL47-like or g60 gene. The attenuated virus is
CC      useful as a vaccine against ILTV. A foreign gene encoding an
CC      antigen e.g. from another avian virus can be inserted into the US2,
CC      UL47-like, ORF4, gG, g60 or gI gene to provide a multivalent
CC      vaccine for chickens and other poultry. Deletion of the gG or gI
CC      gene provides as a negative marker to distinguish vaccinated from
CC      infected animals.
XX      SQ      Sequence 18912 BP; 4416 A; 5261 C; 5251 G; 3984 T; 0 other;

Query Match      100.0%; Score 1305; DB 17; Length 18912;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1      ATGCACCGTCTCATCTCAGACGGGCACTCGCGTTACTACGCGAAAGGAGAGGTGCTTAAC 60
Db      11098      atgcaccgtcctcatctcagacgagcactcgcgttactacgcgaaagagagtgcttaac 11157
QY      61      AAACACATGATGCGGTGGAAGAAACGGTGTGCTGCTCAGCGCAGCTGTATTCACTCTTTTC 120
Db      11158      aaacacatgatgtgcgttggaagaaacggtgtcgtcagcgagcagctgtattcaactcttttc 11217
QY      121      TGGACTTGTGTACAGATTATGCGGGAGCATATCTGCTTTGTAGCAACGCTATGACCGC 180
Db      11218      tggacttgtgtcagattatgctgggagacatatctgtttgtacgaacgctatggaaccg 11277
QY      181      CATTTATTTTGAGGAATGCTTTTGGACTATCGTACTGCTTCTCTCGCTAGCCAG 240
Db      11278      catttattttgaggaatgcttttggactatcgtactgcttctcctcgttagccag 11337
QY      241      AGCACCGCGCGGTACGACTACATTTTAAAGCGGTGCGGCTGAGCGCGCTAACC 300
Db      11338      agcacgcgcgcgtcacgtaagactacatttaggcgcgtcgcgcgtcgcgcgtaacc 11397
```

```
QY      301      ATACCGCGGCTTGCCCGGTATATACAGATACCTCACTAGGGTATCAAGAGGCTGCGACGTT 360
Db      11398      ataccgcggttgcccccgtatataacagatacctcactaggtatcaagagcgtgcgcgtc 11457
QY      361      GTCGAGCTCAACCCGATTTCTAAGCGTGAGCAGACATGATATCGCGGCCCAAGAAAAAGAG 420
Db      11458      gtcgagctcaacccgatttctaacgttgacgacatgatatcgcgcgccaagaaaaagag 11517
QY      421      AAGGGGGGCCCTTTTCAGAGCGCTCCGTCGCTGCTGTTCTACGTGATTAAGGGCGACGACGGC 480
Db      11518      aaggggggcccttttcagagcgctccgctcgtcgtctgttctaaggttaagggcgacgagc 11577
QY      481      GAGGACGAAGTACTGTCTCAATCTATAGAAAAGAGTACAGGAATGTGGCAGCTACAACTG 540
Db      11578      gaggacaagtaactgtccaatctatagaaaagagtaacaggaatgtygcagctacaaactg 11637
QY      541      CTATCTGAATGCGCGGTTCAATCTGACAGAGATGTGGCAGTGAATGTTCTCTAGCACC 600
Db      11638      ctatctgaatgcgccgttcaatctgcacagatgtggtgcagtgactatgttccctagcacc 11697
QY      601      CTTGTATCGCGGAATGCGCGGAGCTGACTATATCTCCCCCACTGCTGCGCTCTGTGCG 660
Db      11698      ctgtatcgcgaatatgycgcygactgactatatttcccccaactgtcgcctcctcgtgc 11757
QY      661      CAATACTTGTCTGACCCCTGAATAATCGGGAGATTGCGCAACAGCTCTCGTAACCTTAGAA 720
Db      11758      caatacttgtaaccctgaataatcgggagatttgcgcaaacagctcgttaactctagaa 11817
QY      721      GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACCTTTTACCGGTGAATGCTGG 780
Db      11818      gttaacgatcgctgtttaagatcgggtcgcagcttaacttttacgcgtgaatgctg 11877
QY      781      ACAACAGAAAGTATCAGACTGGAATTTCAAGGCGAAGACCTTTATCCGATCGCAGACACC 840
Db      11878      acaacagaagatatacagactggaatttcaagcggaacaccttataccgatacgcagacacc 11937
QY      841      AATACACGACACGCGGACGACGCTATATCGGGATACGAAGATATTCTGACGCGTGAAT 900
Db      11938      aatacacgacacgcgcgagcagctatatacgggatacgaagatatctgcagcgctggaat 11997
QY      901      AATTGCTGAGGAAAAAGATCTTAGCGCGCCAGACCCCTCGTCCAGATAGCGTCCCGCAA 960
Db      11998      aattgctgagaaaaagatcttagcgcgcagaccctcgtccagatagcgtcccgcaa 12057
QY      961      GAAATTCCCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGACGCGAGAAAGCGGAA 1020
Db      12058      gaatttcccgctgttaaccagaagaaagcggaagggcgaccccgagcagaaagcagcga 12117
QY      1021      AAGAAGCCCCCTCCAGAAGACTCGGAGGACGACATGCAAGCAGAGGCTTCTGAGAAAT 1080
Db      12118      aagaagccccctccagaagactcggagagcagacatgcaagcagagcagatgtatccaaac 12177
QY      1081      CCTGCCGCCCTCCCGAAGACGACGAAGTCCCCGAGACACCGAGCAGCAGATGATCCAAC 1140
Db      12178      cctgccgccctcccggaagacgaagtcgcccgagacacgagcagatgtatccaaac 12237
QY      1141      TCGGATCCTGACTATTACATGACATGCCCCGCGGTGATCCCGGTGAGAGAGACTACTAAA 1200
Db      12238      tcggatcctgactattacaatgacatgccccgcgtgtatcccgltggaagagactactaaa 12297
QY      1201      AGTTCTAATGCGGCTTCATGCCCCATATTGCGGGCGGCTTGCTGAGCCGCGGCTGCGGCTC 1260
Db      12298      agttctaatgccgcttcacatgccatatcgcgcgttcgttagccctgcgcgctgcgcctc 12357
QY      1261      GTGGGCTACTGTTTGGAGCATGTAATAATGCGCGGTAGCTAA 1305
Db      12358      gtgggctactgtttggagcatgtataaatgcgcgctagctaa 12402
```

RESULT 9
AAQ22986
ID AAQ22986 standard; DNA: 3502 BP.
XX

AC AAQ22986;
XX
DT 02-NOV-1992 (first entry)
XX
DE Sequence of ILTV gp60 gene.
XX
KM Subunit vaccine; immunogen; glycoprotein; promoter; ss.
XX
OS Infectious laryngotracheitis virus.
XX
FH Key
FT TATA_signal Location/Qualifiers
FT 294..299
FT /*tag= a
FT 339..3326
FT /*tag= b
FT 1692..1712
FT /*tag= c
FT 1755..1775
FT /*tag= d
FT 1824..1841
FT /*tag= e
FT 1857..1871
FT /*tag= f
FT 1923..1943
FT /*tag= g
FT 1986..2006
FT /*tag= h
FT 2019..2039
FT /*tag= i
FT 2115..2129
FT /*tag= j
FT 2178..2198
FT /*tag= k
FT 2211..2228
FT /*tag= l
FT 2241..2261
FT /*tag= m
FT 2304..2324
FT /*tag= n
FT 2334..2351
FT /*tag= o
XX
PN W09203554-A.
XX
PD 05-MAR-1992.
XX
PF 23-AUG-1991; 91WO-AU00383.
XX
PR 24-AUG-1990; 90AU-0001937.
XX
PA (WEBSTER-) WEBSTER A PTY LTD.
XX
PI Sheppard MG, Prideaux C, Johnson M, Fahey KJ, York JJ,
PI Kongsuwan K;
XX
DR WPI; 1992-096898/12.
DR P-PSDB; AAR22234.
XX
XX
PT Vaccines against ILTV for use in chickens - comprises
PT non-infectious sub-units or live recombinant viruses and may be
PT administered by aerosol
XX
PS Claim 14; Fig 6; 122pp; English.
XX
CC The inventors claim a non-infectious subunit vaccine for use against
CC ILTV which comprises a glycoprotein of ILTV. They also claim DNA
CC encoding the vaccine, a synthetic polypeptide displaying the
CC antigenicity of the 205K complex or 60K ILTV glycoprotein and a
CC recombinant ILTV with heterologous DNA inserted into a non-essential
CC region of the genome; and a recombinant DNA mol. comprising an ILTV
CC promoter region operatively linked to a heterologous DNA sequence.
CC The promoter region is the ILTV gp60 promoter, gp205 (gpB) promoter
CC or the ORF3 promoter region. KpnK/ORF3 is located 5' of the gp60

CC gene. The deduced AA sequence of the gp60 gene has 19 hydrophobic
CC AA residues at the N-terminus which may correspond to the signal
CC sequence. A second region of hydrophobic AAs (posn. 960 to 989) at
CC the C-terminus could function as a transmembrane anchor sequence.
CC There are nine potential N-linked glycosylation sites on the ILTV
CC gp60 protein. One of these sites at residue 677, may not be active
CC due to the presence of a proline residue within the N-X-S/T signal.
CC There are repeated sequences within the ILTV gp60 coding region.
CC They are not perfectly conserved but several copies had diverged so
CC that only 4 AAs out of 7 were conserved in all 13 copies of the
CC repeats. There is no significant homology between the protein and
CC any sequenced herpesvirus proteins in the database called PIR.
XX
SQ Sequence 3502 BP; 845 A; 969 C; 920 G; 768 T; 0 other;

Query Match 27.7%; Score 361; DB 13; Length 3502;
Best Local Similarity 98.6%; Pred. No. 1.6e-104;
Matches 364; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACCGCAGCTCGCTTACTACGCAAGAGAGGTGCTTAAC 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db atgcaccgtccctcatctcagacgcgcaactcgcgttactacgcgaagagaggtgttaac 3193

QY 61 AAACACATGATTCGGGTGGAAGACGGTGTCTGCTCAGCGCAGCTGTATTCACTCTTTC 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db gaacacatggaattgcggtggaacacggtgtcgtcagcgcaagtgattcacctttc 3253

QY 121 TGGACTTGTGTCAGGATTATGCGGAGCATATCTGCTTGTAGCGCAACGCTATGACCGC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db tggacttgtgtcaggattatgctggaagacatattcgttgtacgcaacgctatgacgcgc 3313

QY 181 CATTTATTTTGAGGAATGCTTTTGGACTATGCTACTGCTTCTTCCTTCGCTAGCCAG 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db cattatttttgaggaatgcttttggactatcgtactgcttcttccttcgtaaccag 3373

QY 241 AGCACC GCCCGCTCAGCTACGACTATTTTGAAGCGGTGCGCGCTCGACGCTAACC 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db agcaccgcgcgcgtcagctacgactacattttaggcgcgtcgcgcgtcgaagcgttaacc 3433

QY 301 ATACCGGCGGTTGGCCGTTAATACAGATACCTACTAGGGTATCAAGAGGCTGCGACGTT 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ataccgcgcggttgcccgataacagatacctcaactaggtatcaagaggtcgcagcct 3493

QY 361 GTCGAGCTC 369
Db ||||||||
Db gtcgagctc 3502

RESULT 10
AAT44399
ID AAT44399 standard; DNA; 42 BP.
XX
AC AAT44399;
XX
DT 03-JUN-1997 (first entry)
XX
DE Homology vector 562-61.1F junction B.
XX
XX
KW ILTV; vaccine; homology vector 562-61.1F; attenuation; poultry;
KW avian infectious bronchitis virus; Newcastle disease virus;
KW infectious bursal disease virus of chickens;
KW Marek's disease virus; herpesvirus; glycoprotein gI; ss.
XX
OS Chimeric infectious laryngotracheitis virus.
XX
FH Key
FT misc_RNA Location/Qualifiers
FT 1..24
FT /*tag= a
FT /*label= ILTV
FT /note= "3' end of approx. 1619 bp Asp718I-XbaI
FT subfragment of the ILTV 8.0 kb Asp718I
FT fragment"

FT	misc_RNA	19..42	
FT		/*tag= b	
FT		/label= ILTV	
FT		/note= "5' end of approx. 691 bp XbaI-XhoI fragment	
FT		of ILTV, generated by PCR"	
XX			
PN	WO9629396-A1.		
XX			
PD	26-SEP-1996.		
XX			
PF	21-MAR-1996;	96WO-US03916.	
XX			
PR	06-JUN-1995;	95US-0468190.	
PR	23-MAR-1995;	95US-0410121.	
XX			
PA	(SYTR) SYNTRO CORP.		
XX			
PI	Cochran MD, Wild MA;		
XX			
DR	WPI; 1996-443172/44.		
XX			
PT	Recombinant infectious laryngotracheitis virus with deletion in the		
PT	glycoprotein G, gi or US2 gene, etc. - useful for vaccines against		
PT	infectious laryngotracheitis in poultry		
XX			
PS	Example 4; Fig 7B; 216pp; English.		
XX			
CC	Junction sequences (AAT44398-403) are provided of the DNA fragments		
CC	used to construct homology vector 562-61.1F, a plasmid designed for		
CC	the purpose of deleting a 983 bp portion of the glycoprotein gi		
CC	gene from the unique short region (see also AAT44384) of infectious		
CC	laryngotracheitis virus (ILTV), and replacing it with uida DNA.		
CC	The vector incorporates a screenable marker, the E. coli uida gene,		
CC	flanked by ILTV DNA. The homology vector is used to construct		
CC	attenuated ILTV S-ILT-011, which is useful as a killed vaccine to		
CC	protect chickens from ILT disease and as a negative marker to		
CC	discriminate vaccinated from infected animals. It is also used to		
CC	construct S-ILT-013, which also has a deletion in the gg gene.		
XX			
SO	Sequence 42 BP; 12 A; 11 C; 8 G; 11 T; 0 other;		
Query Match 3.2%; Score 42; DB 17; Length 42;			
Best Local Similarity 100.0%; Pred. No. 0.0019;			
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	695 CGCAACAGCTCTCGTAACTCTAGAGTTAAGATCGCTGTT	736	
Db	1 cgcaacagctctcgtaactctagaagttaacgacgtgtt	42	
RESULT 11			
ID	AAQ21833		
AC	AAQ21833 standard; DNA; 390 BP.		
XX			
AC	AAQ21833;		
XX			
DT	08-JUN-1992 (first entry)		
XX			
DE	Randomising oligonucleotide used in SPERT mRNA prepn.		
XX			
KW	Systematic polypeptide evolution by reverse translation; SPERT;		
KW	ligand binding; ss.		
XX			
OS	Synthetic.		
XX			
PN	WO9202536-A.		
XX			
PD	20-FEB-1992.		
XX			
PF	01-AUG-1991;	91WO-US05463.	
XX			
PR	02-AUG-1990;	90US-0561968.	

```
XX PA      (COLS ) UNIV OF COLORADO.  
XX PI      Gold L,   Turk C;  
XX DR      WPI; 1992-080018/10.  
XX PT      New method of systematic polypeptide evolution by reverse  
PT translation - by linking each polypeptide in sample mixt. to  
PT individualised mRNA allowing further synthesis of selected  
PT polypeptide(s)  
PS Example; Page 55; 102pp; English.  
XX  
CC The sequence is that of an example randomising oligonucleotide which  
CC is used in the prepn. of mRNA encoding candidate polypeptides for the  
CC method of systematic polypeptide evolution by reverse translation  
CC (SPERT). The method provides a rapid way of isolating and identifying  
CC polypeptide ligands which bind to target mols. The polypeptide ligands  
CC can be used in e.g. assay methods, diagnostic procedures, cell sorting,  
CC as activators or inhibitors of target mol. function, as probes, as  
CC sequestering agents, drug delivery vehicles, modifiers of hormone  
CC action and as catalysts. See also AAQ21830-Q21832.  
CC XX  
SQ Sequence 390 BP; 125 A; 126 C; 133 G; 6 T; 0 other;  
  
Query Match          3.2%; Score 41.6; DB 13; Length 390;  
Best Local Similarity 51.6% ; Pred. No. 0.0086;  
Matches    95; Conservative     0; Mismatches    89; Indels       0; Gaps        0;  
  
OY 1003 GACGCAGAAAGCAGCGAAAAGAAGGCCCTTCAGAGACTCGGAGAGCACATGCAGCA 1062  
Db   ||| | | | | | | | | | | | | | | | | |  
DB 201 gacgacgcagcagcacgcagcgaccgacgcagcagcagcagcagcagcagcagcagcagc 260  
  
OY 1063 GAGGCTTCTGGAGAAAAATCCTCCCCCCCCTCCGAAGACGACGAGTCCCCGAGGACC 1122  
Db   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 261 gagcgacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 320  
  
OY 1123 GAGCAGCATGATCCAAACTCGGATCCTGACTATTACAATGACATGCCCGGTGATCCCG 1182  
Db   || |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 321 gagcgacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcatgttcg 380  
  
OY 1183 GTGG 1186  
Db   || |  
DB 381 atgg 384  
  
RESULT 12  
AAQ36859  
ID AAQ36859 standard; DNA; 390 BP.  
XX AC AAQ36859;  
XX DT 22-JUN-1993 (first entry)  
DE PCR primer for 5' fixed sequence contg. T7 promoter and RBS.  
XX KW Systematic peptide evolution by reverse translation; SPERT; ligand;  
KW specific; inhibitors; probes; assay; cell sorting; ss.  
XX OS Synthetic.  
OS WO9303172-A.  
PN PN  
PD PD 18-FEB-1993.  
PF PF 31-JAN-1992; 92WO-US00801.  
PR PR 01-AUG-1991; 91US-0739055.  
PA (UYRE-) UNIV RES CORP.
```


OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	2.9%;	Score 38.4;	DB 21;	Length 1761;
Best Local Similarity	46.1%;	Pred. No. 0.21;		
Matches 129; Conservative	0;	Mismatches 151;	Indels 0;	Gaps 0;

[illegible]

RESULT	15
ABL11609	
ID	ABL11609 standard; cDNA; 1232 BP.
XX	
AC	ABL11609;

XX	26-MAR-2002 (first entry)
DT	
XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 29309.
DE	
XX	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
KW	
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
DR	P-PSDB; ABB67506.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Claim 1; SEQ ID NO 29309; 21pp + Sequence Listing; English.

AA
PS Claim 1; SEQ ID NO 29309; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1232 BP; 358 A; 319 C; 336 G; 219 T; 0 other;

	Query Match	2.9%;	Score 38.2;	DB 23;	Length 1232;
	Best Local Similarity	55.7%;	Pred. No. 0.2;		
	Matches	73;	Conservative	0;	Mismatches 58; Indels 0; Gaps
QY	1003	GACGCAGAAAGCAGCGAAAAAAGGCCCCCTCCAGAAGACTCGGAGGACGACATGCAGCA	1062		
Dd	183	gcgcgccctccagcagcaaaaaaacgccaacaagaaggysgaagccagtgtgatattaaagaa	242		
QY	1063	GAGGCTTCTGGAGAAAATCTGCCGCCCTCCCGGAAGACGACGAAGTCCCAGAGCACCC	1122		
Dd	243	gaggtatgtatgacgcagcatgtgagggagaaatgtacgcagcagcagcaagaagaagatgtatgac	302		
QY	1123	GAGCACCATGA	1133		
Dd	303	gacgcagcacga	313		

Search completed: September 29, 2002, 01:39:26
Job time: 5143 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2002, 00:11:43 ; Search time 65.96 Seconds
(without alignments)
4859.790 Million cell updates/sec

Title: US-09-994-064-10
Perfect score: 1305
Sequence: 1 ATGCACCGCTCCTCATCTCAG.....TAAATGCGCGGTAGCTAA 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	1305	100.0	1305	US-08-484-575A-19	Sequence 19, Appl
2	1305	100.0	1305	US-08-477-459-19	Sequence 19, Appl
3	1305	100.0	1305	US-08-479-869-19	Sequence 19, Appl
4	1305	100.0	1305	US-08-486-414-19	Sequence 19, Appl
5	1305	100.0	1305	PCT-US94-01826A-19	Sequence 19, Appl
6	1305	100.0	1305	PCT-US94-02252A-19	Sequence 19, Appl
7	1305	100.0	1305	PCT-US96-03916-10	Sequence 10, Appl
8	1305	100.0	13473	PCT-US96-03916-1	Sequence 1, Appl
9	1305	100.0	18912	PCT-US96-03916-59	Sequence 59, Appl
10	44.4	3.4	7218	US-08-232-463-14	Sequence 14, Appl
11	42	3.2	42	PCT-US96-03916-33	Sequence 33, Appl
12	41.6	3.2	390	US-09-197-649-7	Sequence 7, Appl
13	37.4	2.9	2277	US-08-676-967-2	Sequence 2, Appl
14	37.4	2.9	2277	US-08-676-974-2	Sequence 2, Appl
15	37.4	2.9	2277	US-09-098-487-2	Sequence 2, Appl
16	37	2.8	1288	US-08-440-856A-9	Sequence 9, Appl
17	36.2	2.8	2943	US-08-042-747A-7	Sequence 7, Appl
18	35.4	2.7	3489	US-08-728-323A-1	Sequence 1, Appl
19	35.4	2.7	32207	US-08-770-379-20	Sequence 20, Appl
20	35.4	2.7	32207	US-08-757-669A-20	Sequence 20, Appl
21	35.4	2.7	32207	US-09-230-371A-20	Sequence 20, Appl
22	35.2	2.7	150	US-07-829-461A-8	Sequence 8, Appl
23	34.8	2.7	1187	US-08-440-856A-2	Sequence 2, Appl
24	34.6	2.7	1191	US-09-386-588A-1	Sequence 1, Appl
25	34.4	2.6	68750	US-09-335-409-1	Sequence 1, Appl
26	34.4	2.6	68750	US-09-568-102-1	Sequence 1, Appl
27	34.4	2.6	68750	US-09-567-969-1	Sequence 1, Appl

28	34.4	2.6	68750	4	US-09-568-480-1	Sequence 1, Appl
29	34.4	2.6	68750	4	US-09-568-486-1	Sequence 1, Appl
30	34.4	2.6	68750	4	US-09-568-472-1	Sequence 1, Appl
31	34.2	2.6	799	4	US-08-998-416-420	Sequence 420, App
32	34.2	2.6	5063	1	US-08-185-432-1	Sequence 1, Appl
33	33.4	2.6	1100	3	US-09-248-335-53	Sequence 53, Appl
34	33.2	2.5	4403765	4	US-09-103-840A-2	Sequence 2, Appl
35	32.8	2.5	2984	5	PCT-US93-00893-1	Sequence 1, Appl
36	32.8	2.5	2984	5	PCT-US93-00893-2	Sequence 2, Appl
37	32.8	2.5	3777	3	US-09-121-321-15	Sequence 15, Appl
38	32.8	2.5	3777	4	US-08-933-803A-15	Sequence 15, Appl
39	32.4	2.5	1476	4	US-09-434-288-12	Sequence 12, Appl
40	32.4	2.5	28804	2	US-08-592-874-1	Sequence 1, Appl
41	32.4	2.5	28804	3	US-09-096-942-2	Sequence 2, Appl
42	32.4	2.5	28804	3	US-09-096-867-2	Sequence 2, Appl
43	32.2	2.5	2745	1	US-08-363-255-1	Sequence 1, Appl
44	32.2	2.5	2745	1	US-08-363-255-13	Sequence 13, Appl
45	32	2.5	4257	2	US-08-690-473-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-484-575A-19
: Sequence 19, Application US/08484575A
: Patent No. 5925358
: GENERAL INFORMATION:
: APPLICANT: Mark D. Cochran and David E. Junker
: TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: John P. White
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,575A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: White Esq, John P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)278-0450
: TELEFAX: (212)391-0525
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1305 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1305
: US-08-484-575A-19

Query Match 100.0%; Score 1305; DB 2; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGCTCCTCATCTCAGACGCGACTCGGTTACTACGGAAGAGAGGTGCTTAAAC 60
|||||

Db 1 ATGCACCGCTCCTCATCTCAGACGGCAGCTCGCGTTACTACCGGAAAGAGAGGTGCTTAAC 60
QY 61 AAACACATGATTTGCGGTGGAACACGGTGTCTGTCTCAGCGCCAGCTGTATTCTCTTTTC 120
Db 61 AAACACATGATTTGCGGTGGAACACGGTGTCTGTCTCAGCGCCAGCTGTATTCTCTTTTC 120
QY 121 TGGACTTGTGTAGAGATTAATGCGGGAGCATATCTGCTTTGTACGCAACGCTATGGACCGC 180
Db 121 TGGACTTGTGTAGAGATTAATGCGGGAGCATATCTGCTTTGTACGCAACGCTATGGACCGC 180
QY 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCTGCTTCTCTCTCTGCTAGCCAG 240
Db 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCTGCTTCTCTCTCTGCTAGCCAG 240
QY 241 AGCACCGCGCGCTCACTACGACTACATTTTAGCGCCGCTCGCGCTCGACGCGCTAAC 300
Db 241 AGCACCGCGCGCTCACTACGACTACATTTTAGCGCCGCTCGCGCTCGACGCGCTAAC 300
QY 301 ATACCGCGGTTGCGCCGTATACAGATATCTCACTAGGGTATCAAGAGGCTGCGAGCTT 360
Db 301 ATACCGCGGTTGCGCCGTATACAGATATCTCACTAGGGTATCAAGAGGCTGCGAGCTT 360
QY 361 GTCGAGCTCAACCCGATTTCTAACGTGACAGACATGATATCGCGGCCAAAGAAAAAGAG 420
Db 361 GTCGAGCTCAACCCGATTTCTAACGTGACAGACATGATATCGCGGCCAAAGAAAAAGAG 420
QY 421 AAGGGGGCCCTTTTGAAGCCCTCGCTGTCTGTCTACGTGATTAAGGGCGACGAGCGC 480
Db 421 AAGGGGGCCCTTTTGAAGCCCTCGCTGTCTGTCTACGTGATTAAGGGCGACGAGCGC 480
QY 481 GAGGACAAGTACTGTCCATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACACTG 540
Db 481 GAGGACAAGTACTGTCCATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACACTG 540
QY 541 CTATCTGAATGCGCCGTTCAATCTGCACAGATGTGGGAGTGGAGTATGTTCCCTAGACC 600
Db 541 CTATCTGAATGCGCCGTTCAATCTGCACAGATGTGGGAGTGGAGTATGTTCCCTAGACC 600
QY 601 CTTGTATCGCGAAATGGCGCGGAGTGAATATATCTCTCCCACTGCTGCGCTCTCTGCG 660
Db 601 CTTGTATCGCGAAATGGCGCGGAGTGAATATATCTCTCCCACTGCTGCGCTCTCTGCG 660
QY 661 CAATACTGTGCTGACCCCTGAAAAATCGGGAGATTTGGCGAAACAGCTCTCGTACTCTAGAA 720
Db 661 CAATACTGTGCTGACCCCTGAAAAATCGGGAGATTTGGCGAAACAGCTCTCGTACTCTAGAA 720
QY 721 GTTAACGATCGCTGTTTAAAGATCGGGTCCGACGTTAACTTTTACCCTCGAAATGCTGG 780
Db 721 GTTAACGATCGCTGTTTAAAGATCGGGTCCGACGTTAACTTTTACCCTCGAAATGCTGG 780
QY 781 ACAACAGAACAGTATCAGACTGGATTTCAGGGCGGAACACCTTTATCCGATCGACAGACC 840
Db 781 ACAACAGAACAGTATCAGACTGGATTTCAGGGCGGAACACCTTTATCCGATCGACAGACC 840
QY 841 AATACAGACACGCGGACGACGTATATCGGGGATACGAAGATATTCGACGCGTGAAT 900
Db 841 AATACAGACACGCGGACGACGTATATCGGGGATACGAAGATATTCGACGCGTGAAT 900
QY 901 AATTGCTGAGGAAAAAAGATCTAGCGCGCGCAGACCTCTCTCAGATAGCGTCCCGCAA 960
Db 901 AATTGCTGAGGAAAAAAGATCTAGCGCGCGCAGACCTCTCTCAGATAGCGTCCCGCAA 960
QY 961 GAAATTCCTGCTTAACCAAGAAAGCGGAGGCGCACCCCGGACGAGAAAGACGCGAA 1020
Db 961 GAAATTCCTGCTTAACCAAGAAAGCGGAGGCGCACCCCGGACGAGAAAGACGCGAA 1020
QY 1021 AAGAAGGCGCGCTCAGAAAGACTCGAGAGCAGACATGACAGGCGAGGCTTCTGGAGAAAT 1080
Db 1021 AAGAAGGCGCGCTCAGAAAGACTCGAGAGCAGACATGACAGGCGAGGCTTCTGGAGAAAT 1080
QY 1081 CCGCGCGCGCTCCCGAAGACGACGAGTCCCGGAGGACACCGACAGCATGATGCTAAAC 1140
Db 1081 CCGCGCGCGCTCCCGAAGACGACGAGTCCCGGAGGACACCGACAGCATGATGCTAAAC 1140

QY 1141 TCGATCCTGACTATTACATGACATGCCCGCGTGATCCCGGTGAGAGACTACTAAA 1200
Db 1141 TCGATCCTGACTATTACATGACATGCCCGCGTGATCCCGGTGAGAGACTACTAAA 1200
QY 1201 AGTTCTAATGCCGCTCCATGCCCATATTTGCGGGGTTCTGTAGCCTGCGCGTCCGCTC 1260
Db 1201 AGTTCTAATGCCGCTCCATGCCCATATTTGCGGGGTTCTGTAGCCTGCGCGTCCGCTC 1260
QY 1261 GTGGGCTACTGTTTGGAGCATCTAAATGCGCGCTAGCTAA 1305
Db 1261 GTGGGCTACTGTTTGGAGCATCTAAATGCGCGCTAGCTAA 1305

RESULT 2
US-08-477-459-19
; Sequence 19, Application US/08477459
; Patent No. 6001369
; GENERAL INFORMATION:
; APPLICANT: Mark D. Cochran
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,459
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
US-08-477-459-19

Query Match 100.0%; Score 1305; DB 3; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCACCGTCTCATCTCAGACGGCAGCTGCTTACTACGGAAGAGAGGTGCTTAAC 60
Db 1 ATGCACCGTCTCATCTCAGACGGCAGCTGCTTACTACGGAAGAGAGGTGCTTAAC 60
QY 61 AAACACATGATTTGCGGTGGAACACGGTGTCTGTCTCAGCGCCAGCTGTATTCTCTTTTC 120
Db 61 AAACACATGATTTGCGGTGGAACACGGTGTCTGTCTCAGCGCCAGCTGTATTCTCTTTTC 120
QY 121 TGGACTTGTGTAGAGATTAATGCGGGAGCATATCTGCTTTGTACGCAACGCTATGGACCGC 180
Db 121 TGGACTTGTGTAGAGATTAATGCGGGAGCATATCTGCTTTGTACGCAACGCTATGGACCGC 180

Db 121 TGGACTTGTCTCAGGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
QY 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTTCTTCTTCGCTAGCCAG 240
Db 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTTCTTCTTCGCTAGCCAG 240
QY 241 AGCACCGCGCCGCTCAGCTACGACTATCTTTAGGCGCGTCCGCGCTCGACGCGCTAAC 300
Db 241 AGCACCGCGCCGCTCAGCTACGACTATCTTTAGGCGCGTCCGCGCTCGACGCGCTAAC 300
QY 301 ATACCGCGGTTGGCCCGTATAACAGATACCTCCTAGGGTATCAAGAGGCTGCGACGTT 360
Db 301 ATACCGCGGTTGGCCCGTATAACAGATACCTCCTAGGGTATCAAGAGGCTGCGACGTT 360
QY 361 GTCAGCTCAACCCGATTTCTTAACGTGACGACATGATATGCGCGGCCAAAGAAAAAGAG 420
Db 361 GTCAGCTCAACCCGATTTCTTAACGTGACGACATGATATGCGCGGCCAAAGAAAAAGAG 420
QY 421 AAGGGGCGCCCTTTCGAGCGCTCGCTGCTGTTCTACGTGATTAAGGCGACGACGCG 480
Db 421 AAGGGGCGCCCTTTCGAGCGCTCGCTGCTGTTCTACGTGATTAAGGCGACGACGCG 480
QY 481 GAGACAAAGTACTGTCCAATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAAC 540
Db 481 GAGACAAAGTACTGTCCAATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAAC 540
QY 541 CTATCTGAATGCGCGCTCAATCTGACACAGATGTGGGCGACGTGATGTTCTTAGCACC 600
Db 541 CTATCTGAATGCGCGCTCAATCTGACACAGATGTGGGCGACGTGATGTTCTTAGCACC 600
QY 601 CTTGTATCGCGAAATGGCGGCGACTGACTATATCTCCCGCACTGCTGCGCTCTGCG 660
Db 601 CTTGTATCGCGAAATGGCGGCGACTGACTATATCTCCCGCACTGCTGCGCTCTGCG 660
QY 661 CAATCTTGTGCTGACCTGAAATCGGGAGATTTGCGCAACAGCTCTCGTACTCTAGAA 720
Db 661 CAATCTTGTGCTGACCTGAAATCGGGAGATTTGCGCAACAGCTCTCGTACTCTAGAA 720
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCGTGAATGCTGG 780
Db 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCGTGAATGCTGG 780
QY 781 ACAACAGAACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGACAGAC 840
Db 781 ACAACAGAACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGACAGAC 840
QY 841 AATACACGACACGCGGACGAGCTATATCGGGGATACGAAGATATTCGACGCGTGAAT 900
Db 841 AATACACGACACGCGGACGAGCTATATCGGGGATACGAAGATATTCGACGCGTGAAT 900
QY 901 AATTGCTGAGAAAAAGAAATCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCGCAA 960
Db 901 AATTGCTGAGAAAAAGAAATCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCGCAA 960
QY 961 GAAATTCCTCGCTGTAAACCAAGAAAGCGGAGGCGCACCCCGGACGAGAAAGCAGCGAA 1020
Db 961 GAAATTCCTCGCTGTAAACCAAGAAAGCGGAGGCGCACCCCGGACGAGAAAGCAGCGAA 1020
QY 1021 AAGAAGGCGCCCTCCAGAACTCGGAGGACGACATGACAGGAGAGGCTTCTGAGAAAT 1080
Db 1021 AAGAAGGCGCCCTCCAGAACTCGGAGGACGACATGACAGGAGAGGCTTCTGAGAAAT 1080
QY 1081 CTTGCGCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGAGCAGATGATCCAAAC 1140
Db 1081 CTTGCGCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGAGCAGATGATCCAAAC 1140
QY 1141 TCGGATCTGACTATTACATGACATGCCCCGCGTATCCCGGTGAGGAGAGACTACTTAA 1200
Db 1141 TCGGATCTGACTATTACATGACATGCCCCGCGTATCCCGGTGAGGAGAGACTACTTAA 1200
QY 1201 AGTTCTAATGCGCTCCATGCCATATTTGCGGCGTTCGTAGCCTGCGCGTCCGCTC 1260
Db 1201 AGTTCTAATGCGCTCCATGCCATATTTGCGGCGTTCGTAGCCTGCGCGTCCGCTC 1260

QY 1261 GTGGGCTACTGCTTTGGAGCATCGTAAATGCGCGCTAGCTAA 1305
Db 1261 GTGGGCTACTGCTTTGGAGCATCGTAAATGCGCGCTAGCTAA 1305
RESULT 3
US-08-479-869-19
; Sequence 19, Application US/08479869
; Patent No. 6123949
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D, Mark D
; TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and
; TITLE OF INVENTION: Uses thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479, 869
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/024, 156
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; US-08-479-869-19
Query Match 100.0%; Score 1305; DB 3; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCGCTCTCATCTCAGACGCACTGCGTTACTACGGAAAGAGAGTGTCTTAC 60
Db 1 ATGACCGCTCTCATCTCAGACGCACTGCGTTACTACGGAAAGAGAGTGTCTTAC 60
QY 61 AAACACATGATTTGCGGTGAAACGCTGCTCAGCGCAGCTGATTCACCTTTTC 120
Db 61 AAACACATGATTTGCGGTGAAACGCTGCTCAGCGCAGCTGATTCACCTTTTC 120
QY 121 TGCATTTGTCTCAGATTTATGCGGAGACATATCTGCTTTTACGCAACGCTATGACCGC 180
Db 121 TGCATTTGTCTCAGATTTATGCGGAGACATATCTGCTTTTACGCAACGCTATGACCGC 180
QY 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTTCTTCTTCGCTAGCCAG 240
Db 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTTCTTCTTCGCTAGCCAG 240

Db 181 CATTTATTTTGAGGAATGCTTTTGGACATACGTACTGCTTCTTCCTTCGCTAGCCAG 240

QY 241 AGCACCGCCCGCTCAGCTAGCTACATTTTAGCCGCTCGCGCGCTCGACGCCGCTAAC 300

Db 241 AGCACCGCCCGCTCAGCTAGCTACATTTTAGCCGCTCGCGCGCTCGACGCCGCTAAC 300

QY 301 ATACCGCGGTTGGCCCGGTATACAGATACCTCACTAGGATATCAAGAGGCTGCGACGTT 360

Db 301 ATACCGCGGTTGGCCCGGTATACAGATACCTCACTAGGATATCAAGAGGCTGCGACGTT 360

QY 361 GTCGAGCTCAACCCGATTTCTAACGTGACGACATGATATCGCGGCCCAAGAAAAAGAG 420

Db 361 GTCGAGCTCAACCCGATTTCTAACGTGACGACATGATATCGCGGCCCAAGAAAAAGAG 420

QY 421 AAGGGGGCCCTTCGAGGCTCCGCTCTGCTTCTACGTGATTAAGGGCGACGACGCG 480

Db 421 AAGGGGGCCCTTCGAGGCTCCGCTCTGCTTCTACGTGATTAAGGGCGACGACGCG 480

QY 481 GAGACACAGTACTGTCCATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAACTG 540

Db 481 GAGACACAGTACTGTCCATCTATAGAAAAAGATACAGGGAATGTGGCGACGTACAACTG 540

QY 541 CTATCTGAATGCGCCGTTCAATCTGCACAGATGTGGGACGTGACTATGTTCTTACGACC 600

Db 541 CTATCTGAATGCGCCGTTCAATCTGCACAGATGTGGGACGTGACTATGTTCTTACGACC 600

QY 601 CTTGTATCGCGAAATGGCGCGGAGCTGACTATATTCCTCCCACTGCTGCGCTCTGCGC 660

Db 601 CTTGTATCGCGAAATGGCGCGGAGCTGACTATATTCCTCCCACTGCTGCGCTCTGCGC 660

QY 661 CAATACTGTGTCACCCCTGAAAAATCGGGAGATTTGCGCAACAGCTCTGTAACCTAGAA 720

Db 661 CAATACTGTGTCACCCCTGAAAAATCGGGAGATTTGCGCAACAGCTCTGTAACCTAGAA 720

QY 721 GTTAAAGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTGAAATGCTGG 780

Db 721 GTTAAAGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTGAAATGCTGG 780

QY 781 ACAACAGAACAGTATCAGACTGGAATTTCAAGGCGAACCCTTATCCGATCGACAGACACC 840

Db 781 ACAACAGAACAGTATCAGACTGGAATTTCAAGGCGAACCCTTATCCGATCGACAGACACC 840

QY 841 AATACACGACACGCGGACGACGTATATCGGGGATACGAAGATATTCGACGCGCTGAAT 900

Db 841 AATACACGACACGCGGACGACGTATATCGGGGATACGAAGATATTCGACGCGCTGAAT 900

QY 901 AATTGCTGAGAAAAAGAAATCCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCCGCAA 960

Db 901 AATTGCTGAGAAAAAGAAATCCTAGCGCGCCAGACCTCTGTCAGATAGCGTCCCGCAA 960

QY 961 GAAATTCGCGCTGTAAACCAAGAAAGCGGAAGGCGCACCCCGGACGCGAAGAACGCGAA 1020

Db 961 GAAATTCGCGCTGTAAACCAAGAAAGCGGAAGGCGCACCCCGGACGCGAAGAACGCGAA 1020

QY 1021 AAGAAAGGCCCCCTCCAGAAAGCTCGAGAGCAGATGCAAGGAGAGAGGCTTCTGGAGAAAT 1080

Db 1021 AAGAAAGGCCCCCTCCAGAAAGCTCGAGAGCAGATGCAAGGAGAGAGGCTTCTGGAGAAAT 1080

QY 1081 CCTGCCGCGCTCCCGGAAAGAGCAGGAATCCCGGAGAGACACCGACGACGATGATCCAAAC 1140

Db 1081 CCTGCCGCGCTCCCGGAAAGAGCAGGAATCCCGGAGAGACACCGACGACGATGATCCAAAC 1140

QY 1141 TCGGATCCTGACTATTACATGACATGCCCCCGGTGATCCCGGTGAGAGAGACTACTAAA 1200

Db 1141 TCGGATCCTGACTATTACATGACATGCCCCCGGTGATCCCGGTGAGAGAGACTACTAAA 1200

QY 1201 AGTCTAATGCGCGTCTCCATGCCCCCATATTCGCGGCGTTCTGAGGAGAGACTACTAAA 1260

Db 1201 AGTCTAATGCGCGTCTCCATGCCCCCATATTCGCGGCGTTCTGAGGAGAGACTACTAAA 1260

QY 1261 GTGGGCTACTGCTTTGGAGCATCGTAAAAATGCGCGGTAGCTAA 1305

Db 1261 GTGGGCTACTGCTTTGGAGCATCGTAAAAATGCGCGGTAGCTAA 1305

RESULT 4

US-08-486-414-19

; Sequence 19, Application US/08486414B

; Patent No. 6136318

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUSES AND USES THEREOF

; FILE REFERENCE: 42771D

; CURRENT APPLICATION NUMBER: US/08/486,414B

; CURRENT FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 19

; LENGTH: 1305

; TYPE: DNA

; ORGANISM: Fowlpox virus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(432)

US-08-486-414-19

Query Match 100.0%; Score 1305; DB 3; Length 1305;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACGGGACCTCGGCTTACTACGCGAAAGAGAGGCTGCTTAAC 60

Db 1 atgcaccgtctcatctcagacgggacctcggcttactacgcgaaagagagggctgctttaac 60

QY 61 AAACACATGATTCGCGGTGGAACCGGTGCTGCTCAGCGCGACGCTGATTCACCTTTTC 120

Db 61 aaacacatgatctcgggtggaacgggtgctgctcagcgcgacgctgattcactctttc 120

QY 121 TGGACTGTGTACGAGTATATGCGGAGCATATCTGCTTGTACGCAACGCTATGACCGC 180

Db 121 tggactgtgtacgagtatatgcggagcatatctgcttgtacgcaacgctatgacggc 180

QY 181 CATTTATTTTGAAGGAATGCTTTTGGACTATGCTACTGCTTCTTCTTCGCTAGCCAG 240

Db 181 catttattttgaggaatgcttttggactatgctactgcttcttcttcgctagccag 240

QY 241 AGCACCGCCCGCTCAGCTAGCACTACATTTTAGCGCGTCCGCGCTCGACGGCGCTAAC 300

Db 241 agcacccgcccgtcagctagcactacattttagcgcgctccgctcgacggcgctaac 300

QY 301 ATACCGCGGCTTGCGCCGCTATTAACAGATACCTCAGTAGGGTATCAAGAGGCTGCGACGT 360

Db 301 ataccgcggttgcgccgctatttaacagatacctcagtagggatatcaagaggctgcgacgt 360

QY 361 GTCGAGCTCAACCCGATTTCTAAGCTGACGACATGATATCGCGGCCCAAGAAAAAGAG 420

Db 361 gtcgagctcaacccgatttctaaagctgacgacatgatatcgcgggccaagaaaaagag 420

QY 421 AAGGGGGCCCTTCGAGGCTCCGCTGCTGCTTCTACGTGATTAAGGGCGACGACGCGC 480

Db 421 aagggggcccttcgagggctccgctgctgcttctacgtgattaaaggcgacgacgcg 480

QY 481 GAGGACAGTACTGTCTCAATCTATAGAAAAAGTACAGGGAATGTGGCGACGCTACAACTG 540

Db 481 gaggacagtactgtctcaatctatagaaaaagtacagggaaatgtggcgacgctacaa 540

QY 541 CTATCTGAATGCGCGCTCAATCTGACAGATGTGGCAGTGGAGCTATGTTCTCTAGCACC 600

Db 541 ctatctgaatgcgcgctcaatctgacagatgtggcagtggtgactatgttcctagcacc 600

QY 601 CTTGTATCGCGAAATGGCGGAGCTGACTATATTCCTCCCACTGCTGCGCTCTGCGC 660

Db 601 ctgtatcgcgaaatggcggagctgactatattctccccaactgctgctctcggc 660

QY 661 CAATACCTGCTGACCCGTGAATAATCGGAGATTGCGCAACAGCTCTCTACTCTAGAA 720
|||||
Db 661 caatactgtgtaccctgaaaaatcggagatttgccgaaacagctctcgtactctagaa 720
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTGGAATGCTGG 780
|||||
Db 721 gttaacgactgctgtttaaagatcgggtcgcagcttaacttttaccctggaatgctgg 780
QY 781 ACACAGACAGATATCAGACTGGATTTCAGGCGGAACACCTTATCCGATCGACAGACAC 840
|||||
Db 781 acaacagacagatcagactgtatccaagcgaaacacttataccgactgcagacaccc 840
QY 841 AATACACGACACGCGACGACGTATATCGGGGATACGAGATATTCGACGCGCTGGAAT 900
|||||
Db 841 aatacagacacgcgacgacgtatatacgggagatacgaagatatctgcagcgctggaat 900
QY 901 AATTGCTGAGAAAAAGAAATCTTAGCGCGCCAGACCCCTCTCCAGATAGCGTCCGCAA 960
|||||
Db 901 aattgctgagaaaaaagaatcttagcgcgccagacccctctccagatagcgctccgcaa 960
QY 961 GAAATTCCTCCGCTGTAAACAGAAAGCGGAAGGCGCACCCCGGACGACGAGAAAGCAG 1020
|||||
Db 961 gaaatcccgctgtaaacaaagaagcggaagggcgaccccgagcgagaaagcagcgaa 1020
QY 1021 AAGAAGGCCCCCTCCAGAGACTCGGAGGACGACATGACGACGAGAGGCTTCTGAGAAAT 1080
|||||
Db 1021 aagaagggccccctccagaagactcggagacgacatgcagagcagaggtctctgagaaat 1080
QY 1081 CCTGCCGCCCTCCCCGAAGACGACGAAGTCCCCGAGGACACCGGACGATGATCCAAAC 1140
|||||
Db 1081 cctgccgccctccccgaagaagcgaagagtcctcccgagacacccgacgacgatgaccaa 1140
QY 1141 TCGGATCCTGACTATTAACATGACATGCCCCCGCGCTGATCCCGGTGAGAGACTACTAA 1200
|||||
Db 1141 tcgatacctgactataacatgacatgccccgctgacccggtgagagactactaaa 1200
QY 1201 AGTCTAATGCCGTCTCCATGCCCATATTCGCGGCGTTCGTAGCCTGCGCGTCCGCTC 1260
|||||
Db 1201 agtctaatagcgtctccatgcccataatcgcgcgttcgtagcctgcgcgctcgcgtc 1260
QY 1261 GTGGGGCTACTGTTTGAAGCATCGTAAATGCGCGCTAGCTAA 1305
|||||
Db 1261 gtggggctactgtttgagcatcgtaaatgcgcgctagctaa 1305

RESULT 5
PCT-US94-01826A-19
; Sequence 19, Application PC/TUS9401826A
; GENERAL INFORMATION:
; APPLICANT: Syntro Corporation, et al.
; TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01826A
; FILING DATE: 28-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525

; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; PCT-US94-01826A-19

Query Match 100.0%; Score 1305; DB 5; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCTCATCTCAGACGGCACTCGCGTTACTACGGAAGAGAGAGTGCTTAAC 60
|||||
Db 1 ATGCACCGTCTCTCATCTCAGACGGCACTCGCGTTACTACGGAAGAGAGAGTGCTTAAC 60
QY 61 AAACACATGGATTGCGGTGGAACACGGGTGCTGCTCAGAGCGGCACTGTATTCTACTCTTTTC 120
|||||
Db 61 AAACACATGGATTGCGGTGGAACACGGGTGCTGCTCAGAGCGGCACTGTATTCTACTCTTTTC 120
QY 121 TGGACTTGTCAGAGATTATGCGGGAGCATATCTGCTTGTAGCGCAACGCTATGACCGC 180
|||||
Db 121 TGGACTTGTCAGAGATTATGCGGGAGCATATCTGCTTGTAGCGCAACGCTATGACCGC 180
QY 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTCTCTTCCCTTCGCTAGCAG 240
|||||
Db 181 CATTTATTTTGGAGAAATGCTTTTGGACTATCGTACTGCTTCTCTTCCCTTCGCTAGCAG 240
QY 241 AGCACCGCCCGCTCACGTACGACTACATTTTAGCCGCGTCCGCTCGACGCGCTAAC 300
|||||
Db 241 AGCACCGCCCGCTCACGTACGACTACATTTTAGCCGCGTCCGCTCGACGCGCTAAC 300
QY 301 ATACCGGCGGTGCGCCGTTAACAAGATACCTCAGAGGATCAAGAGGCTGCAAGTT 360
|||||
Db 301 ATACCGGCGGTGCGCCGTTAACAAGATACCTCAGAGGATCAAGAGGCTGCAAGTT 360
QY 361 GTGAGCTCAACCCGATTTCTAACGTGAGCAGCATGATTCGGCGGCCAAGAAAAGAG 420
|||||
Db 361 GTGAGCTCAACCCGATTTCTAACGTGAGCAGCATGATTCGGCGGCCAAGAAAAGAG 420
QY 421 AAGGGGGCCCTTTCGAGGCGCTCGCTGCTGTTCTACGTGATTAAGGGCGACGAGCGC 480
|||||
Db 421 AAGGGGGCCCTTTCGAGGCGCTCGCTGCTGTTCTACGTGATTAAGGGCGACGAGCGC 480
QY 481 GAGGACAACTACTGTCCAATCTATAGAAAAGAGTACAGGAAATGTGGCGACGTACACTG 540
|||||
Db 481 GAGGACAACTACTGTCCAATCTATAGAAAAGAGTACAGGAAATGTGGCGACGTACACTG 540
QY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGGAGTGAATGTTCCTAGCACC 600
|||||
Db 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGGAGTGAATGTTCCTAGCACC 600
QY 601 CTGTATCGCGAAATGGCGGCGGACTGACTATATTCCTCCCACTGCTGCGCTCTGCGC 660
|||||
Db 601 CTGTATCGCGAAATGGCGGCGGACTGACTATATTCCTCCCACTGCTGCGCTCTGCGC 660
QY 661 CAATACCTGCTGACCCCTGAAAAATCGGGAGATTGCGCAACAGCTCTGTAACCTAGAA 720
|||||
Db 661 CAATACCTGCTGACCCCTGAAAAATCGGGAGATTGCGCAACAGCTCTGTAACCTAGAA 720
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTGGAATGCTGG 780
|||||
Db 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTGGAATGCTGG 780
QY 781 ACAACAGACAGTATCAGACTGGAATTTCAAGGCGAACAACCTTTATCCGATCGCAGACACC 840

Db 781 ACAACAGAACAGATATCAGACTGGATTTCAGAGCCGACACCTTTATCCGATCGCAGACACC 840
QY 841 AATACAGACACCGGAGCAGCAGTATATCGGGGATACGAAGATATCTGCAGCGCTGAAT 900
Db 841 AATACAGACACCGGAGCAGCAGTATATCGGGGATACGAAGATATCTGCAGCGCTGAAT 900
QY 901 AATTGCTGAGGAAAAAGATCTAGCGCGCCAGACCCCTGCTCAGATATAGCGTCCCGCAA 960
Db 901 AATTGCTGAGGAAAAAGATCTAGCGCGCCAGACCCCTGCTCAGATATAGCGTCCCGCAA 960
QY 961 GAAATTCCTGCTGTAAACCAAGAAAGCGGAAAGGCGCACCCCGGACGCAAGAACGAGAA 1020
Db 961 GAAATTCCTGCTGTAAACCAAGAAAGCGGAAAGGCGCACCCCGGACGCAAGAACGAGAA 1020
QY 1021 AAGAAGCCCTCCAGAGACTCGGAGAGACATGACGACAGGAGGCTTCTGAGAAAT 1080
Db 1021 AAGAAGCCCTCCAGAGACTCGGAGAGACATGACGACAGGAGGCTTCTGAGAAAT 1080
QY 1081 CTTGCCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGACGATGATCCAAAC 1140
Db 1081 CTTGCCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGACGATGATCCAAAC 1140
QY 1141 TCGATCTGACTATTACAATGACATGCCCCCGCTGATCCCGGTGAGAGACTACTAA 1200
Db 1141 TCGATCTGACTATTACAATGACATGCCCCCGCTGATCCCGGTGAGAGACTACTAA 1200
QY 1201 AGTCTAATGCGCTCTCCATGCCCATATTCGCGCGCTTCTGAGCTGCGCGCTC 1260
Db 1201 AGTCTAATGCGCTCTCCATGCCCATATTCGCGCGCTTCTGAGCTGCGCGCTC 1260
QY 1261 GTGGGGCTACTGTTTGAAGCATGTAATAATGCGCGCTAGCTAA 1305
Db 1261 GTGGGGCTACTGTTTGAAGCATGTAATAATGCGCGCTAGCTAA 1305

RESULT 6
PCT-US94-02252A-19
; Sequence 19, Application PC/TUS9402252A
; GENERAL INFORMATION:
; APPLICANT: Syntro Corporation, et al.
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02252A
; FILING DATE: 28-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO

; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
PCT-US94-02252A-19

Query Match 100.0%; Score 1305; DB 5; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGCTCTCATCTCAGACGGCAGCTGCTTACTACGGAAGAGAGGTGCTTAAC 60
Db 1 ATGCACCGCTCTCATCTCAGACGGCAGCTGCTTACTACGGAAGAGAGGTGCTTAAC 60
QY 61 AAACACATGATTCGGGTGAAAAACGGTCTGCTCAGCGCAGCTGTATTCACCTTTTC 120
Db 61 AAACACATGATTCGGGTGAAAAACGGTCTGCTCAGCGCAGCTGTATTCACCTTTTC 120
QY 121 TGGACTTGTGTCAGGATTTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCC 180
Db 121 TGGACTTGTGTCAGGATTTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCC 180
QY 181 CATTATTTTGAAGGAATGCTTTTGGACTATGCTACTGCTTTCTCTTCTGCTAGCCAG 240
Db 181 CATTATTTTGAAGGAATGCTTTTGGACTATGCTACTGCTTTCTCTTCTGCTAGCCAG 240
QY 241 AGCACCGCGCGCTCAGTACGACTATTTTAAAGCCGTCGCGCGCTGACGCGCTTAAC 300
Db 241 AGCACCGCGCGCTCAGTACGACTATTTTAAAGCCGTCGCGCGCTGACGCGCTTAAC 300
QY 301 ATACCGGGGTTGGCCCGTATAACAGATACCTCCTAGGCTATCAAGAGGCTGCGAGCT 360
Db 301 ATACCGGGGTTGGCCCGTATAACAGATACCTCCTAGGCTATCAAGAGGCTGCGAGCT 360
QY 361 GTGAGCTCAACCCGATTTCTAACGTGACGACATGATTCGGCGCCAAAGAAAAAGAG 420
Db 361 GTGAGCTCAACCCGATTTCTAACGTGACGACATGATTCGGCGCCAAAGAAAAAGAG 420
QY 421 AAGGGGGCCCTTTGAGGCTCCGCTGCTGCTGCTTACGTGATTAAGGCGCAGCAGCG 480
Db 421 AAGGGGGCCCTTTGAGGCTCCGCTGCTGCTGCTTACGTGATTAAGGCGCAGCAGCG 480
QY 481 GAGGACAGTACTGCTCCATCTATAGAAAAAGATACAGGGAATGTGGCGAGCTACA 540
Db 481 GAGGACAGTACTGCTCCATCTATAGAAAAAGATACAGGGAATGTGGCGAGCTACA 540
QY 541 CTATCTGAATGCGCGCTTCAATCTGACAGAGATGCGCAGTGAATGTTCTTCTAGCACC 600
Db 541 CTATCTGAATGCGCGCTTCAATCTGACAGAGATGCGCAGTGAATGTTCTTCTAGCACC 600
QY 601 CTTGTATCGGAAATGGCGGGAGCTGACTATATCTCCCCACACTGCTGCGCTCTGCG 660
Db 601 CTTGTATCGGAAATGGCGGGAGCTGACTATATCTCCCCACACTGCTGCGCTCTGCG 660
QY 661 CAATACCTTGTGACCTGAAAAATCGGAGATTTGCGCAACAGCTCTGTAACCTAGAA 720
Db 661 CAATACCTTGTGACCTGAAAAATCGGAGATTTGCGCAACAGCTCTGTAACCTAGAA 720
QY 721 GTTAACGATCGCTGTTTAAAGATCGGGTGGAGCTTAACCTTTTACCGTCAATGCTGG 780
Db 721 GTTAACGATCGCTGTTTAAAGATCGGGTGGAGCTTAACCTTTTACCGTCAATGCTGG 780
QY 781 ACAACAGAACAGATCAGACTGGATTTCAAGCGGAACACACCTTTATCCGATCGCAGACACC 840
Db 781 ACAACAGAACAGATCAGACTGGATTTCAAGCGGAACACACCTTTATCCGATCGCAGACACC 840
QY 841 AATACAGACACCGGAGCAGCAGTATATCGGGGATACGAAGATATCTGCAGCGCTGAAT 900
Db 841 AATACAGACACCGGAGCAGCAGTATATCGGGGATACGAAGATATCTGCAGCGCTGAAT 900
QY 901 AATTGCTGAGGAAAAAGATCTAGCGCGCCAGACCCCTGCTCAGATATAGCGTCCCGCAA 960
Db 901 AATTGCTGAGGAAAAAGATCTAGCGCGCCAGACCCCTGCTCAGATATAGCGTCCCGCAA 960

Db 901 AATTGCTGAGGAAAGAATCTGACGGCCAGACCCTGTCAGATAGCGTCCCGCAA 960
QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAGGGCCACACCCCGGACGCAAGAAAGCAGCGAA 1020
Db 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAGGGCCACACCCCGGACGCAAGAAAGCAGCGAA 1020
QY 1021 AAGAAGCCCCCTCCAGAGACTCGGAGGACGACATGACAGGACAGAGGCTTCTGGAGAAAAT 1080
Db 1021 AAGAAGCCCCCTCCAGAGACTCGGAGGACGACATGACAGGACAGAGGCTTCTGGAGAAAAT 1080
QY 1081 CCTGCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGACGACGATGATCCAAAC 1140
Db 1081 CCTGCGCCCTCCCGAAGACGAGAGTCCCGGAGACACCGACGACGATGATCCAAAC 1140
QY 1141 TCGATCCTGACTATTACAATGACATGCCCGCGTGATCCCGGTGGAGAGACTACTAAA 1200
Db 1141 TCGATCCTGACTATTACAATGACATGCCCGCGTGATCCCGGTGGAGAGACTACTAAA 1200
QY 1201 AGTTCATAGCCGCTCCATGCCCATATTCGCGCGTTCGTAGCCTGCGCGGTGCGGCTC 1260
Db 1201 AGTTCATAGCCGCTCCATGCCCATATTCGCGCGTTCGTAGCCTGCGCGGTGCGGCTC 1260
QY 1261 GTGGGCTACTGTTTGAGAGCATGCTAAATGCGCGCGTAGCTAA 1305
Db 1261 GTGGGCTACTGTTTGAGAGCATGCTAAATGCGCGCGTAGCTAA 1305

RESULT 7

PCT-US96-03916-10
; Sequence 10, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS

; LOCATION: 1..1305
PCT-US96-03916-10
Query Match 100.0%; Score 1305; DB 5; Length 1305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACCCGCTCTCATCTCAGACGCGCCTGCTTACTACGCGAAAGAGGTGCTTAAC 60
Db 1 ATGACCCGCTCTCATCTCAGACGCGCCTGCTTACTACGCGAAAGAGGTGCTTAAC 60
QY 61 AACACATGATTCGCGTGAAAAACGGTCTCTCAGCGCAGCTGTATTCACCTTTTC 120
Db 61 AACACATGATTCGCGTGAAAAACGGTCTCTCAGCGCAGCTGTATTCACCTTTTC 120
QY 121 TGGACTGTGTCAAGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
Db 121 TGGACTGTGTCAAGATTATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
QY 181 CATTTATTTTGAAGAAATGCTTTTGAAGTATGCTACTGCTTCTTCTCGTAGCCAG 240
Db 181 CATTTATTTTGAAGAAATGCTTTTGAAGTATGCTACTGCTTCTTCTCGTAGCCAG 240
QY 241 AGCACCCGCCCGCTCACGTACGACTACATTTTAAGCCGTCGCCGCTGACGCGCTAAC 300
Db 241 AGCACCCGCCCGCTCACGTACGACTACATTTTAAGCCGTCGCCGCTGACGCGCTAAC 300
QY 301 ATACCGCGGCTTGCCCGCTATACAGATACCTCCTAGGGTATCAAGAGGCTGCGACGTT 360
Db 301 ATACCGCGGCTTGCCCGCTATACAGATACCTCCTAGGGTATCAAGAGGCTGCGACGTT 360
QY 361 GTCAGCTCAACCCGATTTCTAACGTGACGACATGATATCGCGGCCAAAGAAAAGAG 420
Db 361 GTCAGCTCAACCCGATTTCTAACGTGACGACATGATATCGCGGCCAAAGAAAAGAG 420
QY 421 AAGGGGCCCTTTCGAGGCCCTCCGTCGTGTTCTACGTGATTAAGGGCGACGACGGC 480
Db 421 AAGGGGCCCTTTCGAGGCCCTCCGTCGTGTTCTACGTGATTAAGGGCGACGACGGC 480
QY 481 GAGGACAGTACTGTCCAATCTATAGAAAAGAGTACAGGGAATGTGGCGACGTACACTG 540
Db 481 GAGGACAGTACTGTCCAATCTATAGAAAAGAGTACAGGGAATGTGGCGACGTACACTG 540
QY 541 CTATCTGAATGCGCCGCTCAATCTGACAGATGTGGGACAGTATGTTCTCTAGCACC 600
Db 541 CTATCTGAATGCGCCGCTCAATCTGACAGATGTGGGACAGTATGTTCTCTAGCACC 600
QY 601 CTGTATCGGAAATGCGCGGAGCTGACTATATCTCCCCCACTGCTGCTCTGGC 660
Db 601 CTGTATCGGAAATGCGCGGAGCTGACTATATCTCCCCCACTGCTGCTCTGGC 660
QY 661 CAATACTTCTGACCCCTGAAAATCGGGAGATTGGCGAAACAGCTCTCGTAACTAGAA 720
Db 661 CAATACTTCTGACCCCTGAAAATCGGGAGATTGGCGAAACAGCTCTCGTAACTAGAA 720
QY 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACCTTAACCTTTTACCGTGAATGCTGG 780
Db 721 GTTAACGATCGCTGTTAAAGATCGGGTCGACCTTAACCTTTTACCGTGAATGCTGG 780
QY 781 ACAACAGACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGACACACC 840
Db 781 ACAACAGACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGACACACC 840
QY 841 AATACAGACACGCGGACGACGTATTCGGGGATACGAAGATATCTGACGCGTGAAT 900
Db 841 AATACAGACACGCGGACGACGTATTCGGGGATACGAAGATATCTGACGCGTGAAT 900
QY 901 AATTGCTGAGGAAAGAAATCTTAGCGCGCCAGACCCCTGTCAGATAGCGTCCCGCAA 960
Db 901 AATTGCTGAGGAAAGAAATCTTAGCGCGCCAGACCCCTGTCAGATAGCGTCCCGCAA 960
QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAGGGCGCACCCCGGACGCAAGACGCGAA 1020

Db 961 GAAATTCCTCGTGTAAACAGAGAAGCGGAGAGGCGCCACCCCGAGCGCAGAAAGCAGCGAA 1020
QY 1021 AAGAAGGCCCCCTCCAGAGACTCGGAGAGCAGATGCGAGAGGCTTCTGAGAAAT 1080
Db 1021 AAGAAGGCCCCCTCCAGAGACTCGGAGAGCAGATGCGAGAGGCTTCTGAGAAAT 1080
QY 1081 CCTGCCGCCCCCTCCAGAGAGCAGAGTCCCGGAGAGACACCGAGCAGATGATCCAAAC 1140
Db 1081 CCTGCCGCCCCCTCCAGAGAGCAGAGTCCCGGAGAGACACCGAGCAGATGATCCAAAC 1140
QY 1141 TCGGATCCTGACTATTACATGACATGCCCGCGTGTATCCCGGTGAGAGAGACTATAA 1200
Db 1141 TCGGATCCTGACTATTACATGACATGCCCGCGTGTATCCCGGTGAGAGAGACTATAA 1200
QY 1201 AGTCTAATGCCGCTCTCCATGCCCCATATTCGCGCGTTCGTAGCCTGCGCGTCCGCTC 1260
Db 1201 AGTCTAATGCCGCTCTCCATGCCCCATATTCGCGCGTTCGTAGCCTGCGCGTCCGCTC 1260
QY 1261 GTGGGCTACTGCTTTGGAGCATCGTAAATGCGCGCTAGCTAA 1305
Db 1261 GTGGGCTACTGCTTTGGAGCATCGTAAATGCGCGCTAGCTAA 1305

RESULT 8
PCT-US96-03916-1

; Sequence 1, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; NUMBER OF INVENTIONS: AND USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13473 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1059..2489
; NAME/KEY: CDS
; LOCATION: 2575..4107

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4113..4445
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4609..5487
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5697..8654
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9874..10962
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11159..12658
; NAME/KEY: CDS
; LOCATION: 12665..13447
PCT-US96-03916-1

Query Match 100.0%; Score 1305; DB 5; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCCCTCATCTCAGACGGCACTCGCGTTACTACGCCAAAGAGAGGTGCTTAAC 60
Db 8462 ATGCACCGTCCCTCATCTCAGACGGCACTCGCGTTACTACGCCAAAGAGAGGTGCTTAAC 8521
QY 61 AAACACATGGAATTGCGGTGGAACCGGTGCTGCTCAGCGCCAGCTGATTCACCTCTTTC 120
Db 8522 AAACACATGGAATTGCGGTGGAACCGGTGCTGCTCAGCGCCAGCTGATTCACCTCTTTC 8581
QY 121 TGGACTTGTCAGAGATTATGCGGAGCATATCTGCTTTGACGCAACGCTATGACCGC 180
Db 8582 TGGACTTGTCAGAGATTATGCGGAGCATATCTGCTTTGACGCAACGCTATGACCGC 8641
QY 181 CATTTATTTTGAAGGAATGCTTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAG 240
Db 8642 CATTTATTTTGAAGGAATGCTTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAG 8701
QY 241 AGCACCGCGCGCGTACAGTACATATTTAGCGCGTCCGCGCTGACGCGCTAAC 300
Db 8702 AGCACCGCGCGCGTACAGTACATATTTAGCGCGTCCGCGCTGACGCGCTAAC 8761
QY 301 ATACCGCGGTTGGCCCGTATACAGATACCTACCTAGGCTATCAAGAGCGTGCAGCTT 360
Db 8762 ATACCGCGGTTGGCCCGTATACAGATACCTACCTAGGCTATCAAGAGCGTGCAGCTT 8821
QY 361 GTCGAGCTCAACCGGATTTCTAACGTGAGCAGATGATATCGGCGCAAGAAAGAG 420
Db 8822 GTCGAGCTCAACCGGATTTCTAACGTGAGCAGATGATATCGGCGCAAGAAAGAG 8881
QY 421 AAGGGGGCCCTTTCGAGGCGCTCCGCTGCTGCTTCTACGTGATTAAGGGCGAGCGGC 480
Db 8882 AAGGGGGCCCTTTCGAGGCGCTCCGCTGCTGCTTCTACGTGATTAAGGGCGAGCGGC 8941
QY 481 GAGGACAAGTACTGTCCAATCTATAGAAAGAGTACAGGGAATGTGGCGAGCTACACTG 540
Db 8942 GAGGACAAGTACTGTCCAATCTATAGAAAGAGTACAGGGAATGTGGCGAGCTACACTG 9001
QY 541 CTATCTGAATGCGCGTTCAATCTGCAAGATGTGGCAGTGAATATGTTCTAGCACC 600
Db 9002 CTATCTGAATGCGCGTTCAATCTGCAAGATGTGGCAGTGAATATGTTCTAGCACC 9061
QY 601 CTTGTATCGGGAATGCGCGGAGTACTATATTTCCCGCACTGCTGCGCTCTGCGC 660
Db 9062 CTTGTATCGGGAATGCGCGGAGTACTATATTTCCCGCACTGCTGCGCTCTGCGC 9121
QY 661 CAATACTTGCTGACCCCTGAAATCGGAGATTGCGCAACAGCTCTGTAACCTAGAA 720
Db 9122 CAATACTTGCTGACCCCTGAAATCGGAGATTGCGCAACAGCTCTGTAACCTAGAA 9181

```
QY 721 GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTCGAATGCTGG 780
      |||
Db 9182 GTTAACGATCGCTGTTTAAAGATCGGGTCGACGCTTAACCTTTTACCCTCGAATGCTGG 9241
QY 781 ACAACAGAACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGCAGACACC 840
      |||
Db 9242 ACAACAGAACAGTATCAGACTGATTTCAAGGCGAACACCTTTATCCGATCGCAGACACC 9301
QY 841 AATACACGACACGCGGACGACGATATATCGGGGATACGAAGATATTCGACGCGCTGGAAT 900
      |||
Db 9302 AATACACGACACGCGGACGACGATATATCGGGGATACGAAGATATTCGACGCGCTGGAAT 9361
QY 901 AATTGCTGAGGAAAAAGAAATCCTAGCGCGCCAGACCCCTCTCCAGATAGCGTCCGCAA 960
      |||
Db 9362 AATTGCTGAGGAAAAAGAAATCCTAGCGCGCCAGACCCCTCTCCAGATAGCGTCCGCAA 9421
QY 961 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAAGGCGCACACCCCGGACGCGAAGACGAGAA 1020
      |||
Db 9422 GAAATTCGCCGCTGTAAACCAAGAAAGCGGAAGGCGCACACCCCGGACGCGAAGACGAGAA 9481
QY 1021 AAGAAGGCCCCCTCCAGAAAGACTCGGAGAGACGACATGCAGCGACGAGAGGCTTCTGGAGAAAT 1080
      |||
Db 9482 AAGAAGGCCCCCTCCAGAAAGACTCGGAGAGACGACATGCAGCGACGAGAGGCTTCTGGAGAAAT 9541
QY 1081 CTTGCCCGCCCTCCCGAAGACGAGAAAGTCCCGCGAGACACCGACGACGATGATCCAAAC 1140
      |||
Db 9542 CTTGCCCGCCCTCCCGAAGACGAGAAAGTCCCGCGAGACACCGACGACGATGATCCAAAC 9601
QY 1141 TCGGATCCTGACTATTACAAATGACATGCCCGCGTGTATCCCGGTGAGAGACTACTAAA 1200
      |||
Db 9602 TCGGATCCTGACTATTACAAATGACATGCCCGCGTGTATCCCGGTGAGAGACTACTAAA 9661
QY 1201 AGTTCTAATGCCGTCTCCATGCCCATATTCGCGGCGTTTCGTAGCCTGCGCGTCCGCTC 1260
      |||
Db 9662 AGTTCTAATGCCGTCTCCATGCCCATATTCGCGGCGTTTCGTAGCCTGCGCGTCCGCTC 9721
QY 1261 GTGGGGCTACTGTTTGAGACATCGTAAATGCGCGCGGTAGCTAA 1305
      |||
Db 9722 GTGGGGCTACTGTTTGAGACATCGTAAATGCGCGCGGTAGCTAA 9766

RESULT 9
PCT-US96-03916-59
; Sequence 59, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
```

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18912 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 697..1533
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1900..2784)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2916..3605)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3694..5124
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5210..7081
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7245..8123
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8333..11290
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12510..13598
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13792..15291
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15298..16080
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16129..17013
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (17380..18216)
; OTHER INFORMATION:
PCT-US96-03916-59

Query Match 100.0%; Score 1305; DB 5; Length 18912;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCACCGTCTCATCTCAGACGGCACTGCGGTTACTACGCGAAGAGAGAGGTGCTTAAC 60
      |||
```



```
Db 11098 ATGCACCGCTCTCATCTCAGACGGCAGCTCGCCTTACTAGCGAAAGAGAGGTGCTTAAC 11157
QY 61 AACACATGATGATTCGGGTGGAACCGCTGCTGCTCAGGCGCAGCTGTATTCTACTCTTTTC 120
Db 11158 AAACACATGATGATTCGGGTGGAACCGCTGCTGCTCAGGCGCAGCTGTATTCTACTCTTTTC 11217
QY 121 TGGACTGTGTGAGGATTAATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 180
Db 11218 TGGACTGTGTGAGGATTAATGCGGAGCATATCTGCTTTGTACGCAACGCTATGACCGC 11277
QY 181 CATTTATTTTGGAGGAATGCTTTTGGAGTATCTGCTTCTCTCTCTGCTAGCCAG 240
Db 11278 CATTTATTTTGGAGGAATGCTTTTGGAGTATCTGCTTCTCTCTCTGCTAGCCAG 11337
QY 241 AGCACCGCGCGCTCAGCTACGACTACATTTTAGCGCGCTCGCGCGCTCGACGCGCTAAC 300
Db 11338 AGCACCGCGCGCTCAGCTACGACTACATTTTAGCGCGCTCGCGCGCTCGACGCGCTAAC 11397
QY 301 ATACCGCGCGCTTGGCGCGCTATACAGATACCTCCTAGGCTATCAAGAGGCTGCGAGCTT 360
Db 11398 ATACCGCGCGCTTGGCGCGCTATACAGATACCTCCTAGGCTATCAAGAGGCTGCGAGCTT 11457
QY 361 GTCGAGCTCAACCCGATTTCTTAACGTGGAGCAGCATGATTCGGCGCCAAAGAAAAAGAG 420
Db 11458 GTCGAGCTCAACCCGATTTCTTAACGTGGAGCAGCATGATTCGGCGCCAAAGAAAAAGAG 11517
QY 421 AAGGGGGGGCTTTCGAGGCGCTCCGCTGCTGCTGCTTCTAGTATTAAGGGCGAGCGGC 480
Db 11518 AAGGGGGGGCTTTCGAGGCGCTCCGCTGCTGCTGCTTCTAGTATTAAGGGCGAGCGGC 11577
QY 481 GAGGACAAGTACTGTCCAACTATAGAAAAAGATACAGGGAATGTGGCGAGCTACAACTG 540
Db 11578 GAGGACAAGTACTGTCCAACTATAGAAAAAGATACAGGGAATGTGGCGAGCTACAACTG 11637
QY 541 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGCAGTGAATGTTCTTCTAGCACC 600
Db 11638 CTATCTGAATGCGCGCTTCAATCTGCACAGATGTGGCAGTGAATGTTCTTCTAGCACC 11697
QY 601 CTTGTATCGCGAAATGGCGCGGAGCTGACTATATCTCCCCACTGCTGCGCTCTGCGC 660
Db 11698 CTTGTATCGCGAAATGGCGCGGAGCTGACTATATCTCCCCACTGCTGCGCTCTGCGC 11757
QY 661 CAATACTTGTGACCTGAAAAATCGGAGATTTGGCCAAACAGCTCTCGTAATCTAGAA 720
Db 11758 CAATACTTGTGACCTGAAAAATCGGAGATTTGGCCAAACAGCTCTCGTAATCTAGAA 11817
QY 721 GTTAAAGATCGCTGTTTAAAGATCGGGTCCGAGCTTAATTTTACCCTCGAAATGCTGG 780
Db 11818 GTTAAAGATCGCTGTTTAAAGATCGGGTCCGAGCTTAATTTTACCCTCGAAATGCTGG 11877
QY 781 ACAACAGAACAGTATCAGACTGGATTTCAAGCGCAACACCTTTATCCGATCGCAGACACC 840
Db 11878 ACAACAGAACAGTATCAGACTGGATTTCAAGCGCAACACCTTTATCCGATCGCAGACACC 11937
QY 841 AATACACGACAGCGGAGCAGCAGTATATCGGGGATACGAAGATATTCGACGCTGAAT 900
Db 11938 AATACACGACAGCGGAGCAGCAGTATATCGGGGATACGAAGATATTCGACGCTGAAT 11997
QY 901 AATTGCTGAGGAAAAAAGAAATCTAGCGCGCCAGACCCCTGCTCAGATAGCGTCCGCAA 960
Db 11998 AATTGCTGAGGAAAAAAGAAATCTAGCGCGCCAGACCCCTGCTCAGATAGCGTCCGCAA 12057
QY 961 GAAATTCGCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGGACGAGAAAGCAGCGAA 1020
Db 12058 GAAATTCGCGCTGTAAACCAAGAAAGCGGAAGGGCGCACCCCGGACGAGAAAGCAGCGAA 12117
QY 1021 AAGAAGCGCCCTCCAGAGACTCGAGAGCAGCATGACAGGCGAGAGGCTTCTGGAGAAAT 1080
Db 12118 AAGAAGCGCCCTCCAGAGACTCGAGAGCAGCATGACAGGCGAGAGGCTTCTGGAGAAAT 12177
QY 1081 CTTGCCGCCCTCCCGAAGACGAGCAAGTCCCGAGAGACCGAGCAGCATGATCCAAAC 1140
Db 12178 CTTGCCGCCCTCCCGAAGACGAGCAAGTCCCGAGAGACCGAGCAGCATGATCCAAAC 12237
```

```
QY 1141 TCGGATCTGACTATTATACATGACATGCCCGCGCTGATCCGGTGAGAGAGACTATAA 1200
Db 12238 TCGGATCTGACTATTATACATGACATGCCCGCGCTGATCCGGTGAGAGAGACTATAA 12297
QY 1201 AGTCTAATGCCGTCCCATGCCATATTCGGCGGCTTCGTAGCCTGCGCGTCCGCTC 1260
Db 12298 AGTCTAATGCCGTCCCATGCCATATTCGGCGGCTTCGTAGCCTGCGCGTCCGCTC 12357
QY 1261 GTGGGCTACTGTTTGGAGCATCGTAAATGCGCGCTAGCTAA 1305
Db 12358 GTGGGCTACTGTTTGGAGCATCGTAAATGCGCGCTAGCTAA 12402
```

```
RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-F1s
; US-08-232-463-14
```

```
Query Match 3.4%; Score 44.4; DB 1; Length 7218;
Best Local Similarity 6.6%; Pred. No. 0.0012;
Matches 24; Conservative 186; Mismatches 152; Indels 0; Gaps 0;
```

```
QY 718 GAAGTTAAGATCGCTGTTTAAAGATCGGGTCCGACCTTAATTTTACCCTCGAAATGC 777
Db 1483 GTAATTACCTATCTATGCAAGTAGTTAAAGAGATGAAGAATTTGGTACRRRRRRRRRR 1424
QY 778 TGGACAACAGAACAGTATCAGACTGGATTTTCAAGCGCGAACACCTTTATCCGATCGCAGAC 837
```


[illegible]

```

RESULT 11
PCT-US96-03916-33
; Sequence 33, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
PCT-US96-03916-33

```

	Matches	42;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
OY	695	CGCAACAGCTCTCGTAACTCTAGAGTTAACGATCGCTGTT								736
Db	1	CGCAACAGCTCTCGTAACTCTAGAGTTAACGATCGCTGTT								42

```

RESULT 12
US-09-197-649-7
; Sequence 7, Application US/09197649
; Patent No. 6194550
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Tuerk, Craig
; APPLICANT: Pribnow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/C1-CON
; CURRENT APPLICATION NUMBER: US/09/197,649
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 07/829,461
; EARLIER FILING DATE: 1992-01-31
; EARLIER APPLICATION NUMBER: 07/739,055
; EARLIER FILING DATE: 1991-08-01
; EARLIER APPLICATION NUMBER: 07/561,968
; EARLIER FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

```

```

Query Match          3.28; Score 41.6; DB 4; Length 390;
Best Local Similarity 51.68; Pred. No. 0.0022;
Matches 95; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1003 GACGCAGAAAGCAGCGAAAGAGGCCCCCTCCAGAGACTCGGAGAGCAGCATGCAGGCA 1062
      |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 201 gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 260
      gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac

QY 1063 GAGGCTTCTGAGAAATCTCCGCCCTCCCGAAGACGACGAAGTCCCGAGGACACC 1122
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 261 gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 320
      gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac

QY 1123 GAGCACGATGATCCAAACTCGGATCCTGACTATTACAATGACATGCCCGGTGATCCC 1182
      || |||| | | | | | | | | | | | | | | | | | | | | | |
DB 321 gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 380
      gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac

QY 1183 GTGG 1186
      |||
DB 381 atg 384

RESULT 13
US-08-676-967-2
; Sequence 2, Application US//08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA

```

COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,967
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-967-2

Query Match	2.9%	Score 37.4;	DB 1;	Length 2277;
Best Local Similarity	33.7%;	Pred. No. 0.1;		
Matches	63;	Conservative	34;	Mismatches 90;
				Indels 0;
				Gaps 0;

QY	959	AAGAAATTCCCGCTGTAAACCAAGAAAGCGGAAGGGCCGACCCCCGGAGCGCAGAAGAAGCAGCG	1018
		: : : : : : : : : : :	
Dd	587	ARGAYVANCARWMSGTWNMGNCNATHGGNGARGARAARWSNCAYGARWSMAARCAYCARG	646
QY	1019	AAAAGAAGGCCCTCCAGAAGACTCGAGAGCAGCATGCAGCGCAGAGGCTTCTTGAGAAA	1078
		: : : : : : : : : : : : : :	
Dd	647	ARMNSGTNAARAARAARARGMNGARGARGARGAYATGGARGARGARGAARAAYGAYG	706
QY	1079	ATPCTGCGCGCCCTCCCGAAGACGACGAAGTCCC CGAGGACACCAGCAGCATGATCCAA	1138
		: : : : : : : : : : : : : :	
Dd	707	AYGAYGAYGAYGAYGAYGARGARGAYGNGNTTYGAYGAYGARGAYGARGARGARBARA	766
QY	1139	ACTCGGA	1145
		: :	
Dd	767	AYATHGA	773

RESULT 14
 US-08-676-974-2
 ; Sequence 2, Application US/08676974
 ; Patent No. 5770422
 ; GENERAL INFORMATION:
 ; APPLICANT: COLLINS, KATHLEEN
 ; TITLE OF INVENTION: Human Telomerase
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Science & Technology Law Group
 ; STREET: 268 Bush Street, Suite 3200
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94104
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/676,974
 ; FILING DATE:
 ; CLASSIFICATION: 530

```

ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```

Query Match	2.9%;	Score 37.4;	DB 1;	Length 2277;
Best Local Similarity	33.7%;	Pred. No. 0.1;		
Matches 63;	Conservative 34;	Mismatches 90;	Indels 0;	Gaps 0;

QY	959	AAGAAATTCCCGCTGTAACCAAGAAAGCGGAAGGGCCACCCCCGGACGCAGAAAGCAGCG	1018
		: : : : : : : : : :	
Dd	587	ARGAYACNCARWMSGTWNMSGNATHGGNGARGARAARWENCAYGARWSMAARCAIYCARG	646
QY	1019	AAAAGAAGGCCCTCCAGAAAGACTCGAGAGCAGCATGCAGCGCAGAGGCTTCTGAGAAA	1078
		: : : : : : : : : : :	
Dd	647	ARMMSGTNARAARAARAARGGNMGNGARGARGARAYATGARGARGARARAAYGAYGAYC	706
QY	1079	ATTCGTCCGCCCTCCCGAAGACGACGAAGTCCCCGAGGACACCAGCAGATGATCCA	1138
		: : : : : : : : : : :	
Dd	707	AYGAYGAYGAYGAYGAYGARGARGAYGNGNTTTYGAYGAYGARGAYGARGARGAR	766
QY	1139	ACTCGGA	1145
		: :	
Dd	767	AYATHGA	773

```

15
US-09-098-487-2
; Sequence 2, Application US/09098487
; Patent No. 5917025
;
; GENERAL INFORMATION:
; APPLICANT: COLLINS, Kathleen
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,487
; FILING DATE:
;
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid

```

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
us-09-098-487-2

Query Match 2.9%; Score 37.4; DB 2; Length 2277;
Best Local Similarity 33.7%; Pred. No. 0.1;
Matches 63; Conservative 34; Mismatches 90; Indels 0; Gaps 0;

QY 959 AAGAAATTCCTCCCTGTACCAAGAAAGCGGAGCGCACCCCGGACGAGAAAGCAGCG 1018
|:| | | : | | : | | | | : | | : | | : |
Db 587 ARGAYACNCARWSNNGTNSNGCNATHGNGARGARAARWSNCAYGARWSNAARCAAYCARG 646
QY 1019 AAAAGAAAGCCCTCCAGAGACTCGAGAGACGACATGCAGGCGAGGCTTCTGGAGAAA 1078
|:| : | | : | | | | | | | | : | | |
Db 647 ARWSNGTNAARAARAARGGNMGNGARGARGAYATGARGARGARAAYAYGAYG 706
QY 1079 ATCTGCCGCCCTCCCGAAGAGACGAAGTCCCCGAGACACCGACGACGATGATCCAA 1138
|:| : | | : : | | | | | | : | | | : | | : |
Db 707 AYGAYGAYGAYGAYGARGARGAYGNGTNTTYGAYGAYGARGAYGARGARGARA 766
QY 1139 ACTCGGA 1145
|:| | |
Db 767 AYATHGA 773

Search completed: September 29, 2002, 01:35:16
Job time: 5013 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 21:59:57 : Search time 1809.52 seconds
(without alignments)
9733.811 Million cell updates/sec

Title: US-09-994-064-10
Perfect score: 1305
Sequence: 1 ATGCACCGCTCCTCATCTCAG.....TAAATGCGCGCTAGCTAA 1305

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.4	3.2	958	12	CNS03PCF AL254472 Tetraodon
2	41.2	3.2	669	10	BE776357 MY-14-G-0
3	40.8	3.1	321	10	BG874723 RCI-BT025
4	40.8	3.1	529	12	AQ876396 V99C8 mtN
5	40.6	3.1	317	9	AA942788 EST198287
6	40.2	3.1	212	9	AJ012890 EHEST023R
7	40.2	3.1	865	10	BG345158 HVSMEG001
8	40.2	3.1	688	10	BF521413 EST458869
C 9	40	3.1	915	12	AL225504 Tetraodon
10	40	3.1	1101	12	CNS00KK2 AL077673 Drosophila
C 11	39.6	3.0	558	10	BF825826 MR2-HN003
C 12	39.6	3.0	567	9	AV896003 AV896003
C 13	39.6	3.0	637	9	AV901789 AV901789
C 14	39.6	3.0	713	9	AV897918 AV897918
C 15	39.6	3.0	715	9	AV892667 AV892667
C 16	39.6	3.0	716	9	AV896867 AV896867
17	39.6	3.0	1046	12	AG073741 Pan trogl

18	39.4	3.0	451	9	AA552908	AA552908 nk60b04.s
19	39.4	3.0	530	10	BM165497	BM165497 EST568020
20	39.4	3.0	677	10	BM161314	BM161314 EST563837
21	39.4	3.0	828	10	BG418286	BG418286 HVSMEK002
22	39.2	3.0	1208	9	AL514927	AL514927 AL514927
23	39	3.0	541	10	BI864958	BI864958 ft17b06.Y
24	38.8	3.0	567	10	BE393657	BE393657 601310380
C 25	38.8	3.0	925	12	CNS0091P	AL053013 Drosophila
26	38.6	3.0	581	9	AV675337	AV675337 AV675337
27	38.4	2.9	295	10	BF908072	BF908072 RCI-UT008
28	38.4	2.9	464	10	BE799902	BE799902 601588045
C 29	38.2	2.9	317	9	AW291911	AW291911 UT-H-BI2-
30	38.2	2.9	478	10	BI366211	BI366211 RE51706.5
31	38.2	2.9	485	9	AA202831	AA202831 LD03155.5
32	38.2	2.9	492	10	BE510710	BE510710 946054H11
33	38.2	2.9	508	10	BI229580	BI229580 RE27390.5
34	38.2	2.9	509	10	BE518669	BE518669 946073B12
35	38.2	2.9	565	10	BI368667	BI368667 RE54674.5
36	38.2	2.9	574	9	AI515323	AI515323 LD47085.5
37	38.2	2.9	589	9	AI514749	AI514749 LD46427.5
38	38.2	2.9	616	10	BI374787	BI374787 RE62461.5
39	38.2	2.9	620	10	BI370037	BI370037 RE56406.5
40	38.2	2.9	638	10	BI162447	BI162447 RE01761.5
41	38.2	2.9	655	10	BI214686	BI214686 RE20684.5
42	38.2	2.9	656	10	BI367087	BI367087 RE52761.5
43	38.2	2.9	826	9	AA940987	AA940987 LD23405.5
44	38.2	2.9	972	12	CNS0137H	AL103295 Drosophila
C 45	38.2	2.9				

ALIGNMENTS

RESULT 1	CNS03PCF/C	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES			
1	CNS03PCF	958 bp	DNA linear	GSS 17-MAY-2000	AL254472	GI:7975484	GSS; genome survey sequence.	Tetraodon nigroviridis	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	1 (bases 1 to 958)	Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	Unpublished	2 (bases 1 to 958)	Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence	Unpublished	3 (bases 1 to 958)	Genoscope.	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .
																		location/Qualifiers			
																		1..958			
																		/organism="Tetraodon nigroviridis"			

BASE COUNT	ORIGIN
253 a	204 c 220 g 271 t 10 others

Query Match	3.2%;	Score 41.4;	DB 12;	Length 958;
Best Local Similarity	53.5%;	Pred. No. 1.5;		
Matches	84;	Conservative	1;	Mismatches 72;
				Indels 0.
				Canc 0.

[illegible]

RESULT	2		
LOCUS	BE776357		
DEFINITION	BE776357	669 bp	mRNA
ACCESSION	MY-14-G-07	pinfestansMY	linear
VERSION	BE776357	Phytophthora infestans	CDNA, mRNA sequence.
KEYWORDS	BE776357.1	GI:10230012	
SOURCE	EST.		
ORGANISM	potato late blight agent.		
	Phytophthora infestans		

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE
1	Kamoun, S., Hrabner, P.T., Sobral, B.W.S., Nuss, D. and Govers, F.	Initial assessment of gene diversity for the oomycete pathogen <i>Phytophthora infestans</i> based on expressed sequences	Fungal Genet. Biol. 28 (2), 94-106 (1999)	20056376

COMMENT

Contact: Govers F
Laboratory of Phytopathology
Wageningen University
Binnenhaven 9, P.O.Box 8025, 6700 EE, Wageningen, The Netherlands
Tel: 31 317 483 138
Fax: 31 317 483 412
Email: Francine.Govers@medew.fyto.wau.nl.
Location/Qualifiers

FEATURES

```

/organism="Phytophthora infestans"
/strain="DDR7602, A1 mating type"
/db_xref="taxon:4787"
/clone_lib="PinfestansMY"
/dev_stage="4-week old vegetative, non-sporulating
mycelium in synthetic medium"
/lab_host="E. coli, strain DH5-alpha"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Total
RNA was isolated from mycelium of P. infestans DDR7602
cultured for 4 weeks in synthetic medium. EST clones were
named by their position in the microtiter plate, preceded
by the prefix MY (for mycelial) and the successive number
of the microtiter plate (e.g. MY-06-A-04)."
```

Query Match	3.28;	Score 41.2;	DB 10;	Length 669;
Best Local Similarity	48.08;	Pred. No. 1.5;		
Matches 118;	Conservative 0;	Mismatches 128;	Indels 0;	Gaps 0;

[illegible]

RESULT	3			
LOCUS	321 bp			
DEFINITION	321 bp	mRNA	linear	EST 30-MAY-2001
ACCESSION	RC1-BT0254-120200-016-c08	BT0254	Homo sapiens cDNA,	mRNA sequence.
VERSION	RC1-BT0254-120200-016-c08			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

REFERENCE AUTHORS	TITLE
Mammalia, Eulimnia; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 321)	
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	

COMMENT
Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=RC1-BT0254-120200-016-c08&ts=2000-02-12&tt=1) Seq primer: puc 18 forward High quality sequence start: 32 High quality sequence stop: 87. Location/Qualifiers 1..321

```

/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      104 a      124 c      89 g      4 t
ORIGIN

```


Db 33 TCGGGCAACACCAAGCCCAAGCTGAGCCAGCAATAGCCGATGGCCGACCCTCGAGAAGCAGAGA 92

QY 1013 GCAGCGAAAAGAG-gccccctccagaagactcgagagacgacatgcagagagcttct 1071

Db 93 CCCGCTGAGAGCCCGCGGAGAGCAGAGTAATGAGGCGCTCAGGGTGGCGAGATGCA 152

QY 1072 GGAGAAAATCTGCGCCCTCCCGGAAGACGAGAGTCCCCGAGACACGACGAT 1131

Db 153 GCACCCGCTGACAGTGCACACTGCACCCGCGAGAGAGAGCCCAAGGCCCGCGAGAGAT 212

QY 1132 GATCCAAACTCGGATCCTGACTATTACATGACATGCC 1170

Db 213 GCGCCCAAGCCCAACCTGACTTTATCGAGAGCTTGCCC 251

RESULT 6

LOCUS AJ012890 212 bp mRNA linear EST 27-MAR-2000

DEFINITION EHST023R Entamoeba histolytica HM-1:IMSS Entamoeba histolytica

CDNA, mRNA sequence.

ACCESSION AJ012890

VERSION AJ012890.1 GI:7330436

KEYWORDS EST.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.

REFERENCE 1 (bases 1 to 212)

AUTHORS Willhoeft,U., Buss,H. and Tannich,E.

TITLE Analysis of cDNA expressed sequence tags from entamoeba histolytica: identification of two highly abundant polyadenylated transcripts with no overt open reading frames

JOURNAL protist 150 (1), 61-70 (1999)

MEDLINE 20187114

COMMENT Contact: Tannich E

Bernhard Nocht Institute for Tropical Medicine

Bernhard Nocht Str. 74, 20359 Hamburg, Germany.

FEATURES

source

1. 212

/organism="Entamoeba histolytica"

/strain="HM-1:IMSS; ATCC 30459"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica HM-1:IMSS"

/note="Vector: lambda ZAP"

BASE COUNT 88 a 30 c 55 g 39 t

ORIGIN

Query Match 3.1%; Score 40.2; DB 9; Length 212;

Best Local Similarity 54.4%; Pred.No.1.7;

Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1003 GACGCAAGACGACGAAAAGAGGCCCTCCAGAAGACTCGAGAGCAGATGCAGCA 1062

Db 59 GACTTAGAAACAACACTACACAACGAATTCGAAGACGACGATGACGACTATGACTTA 118

QY 1063 GAGGCTCTGGAGAAATCTGCGCCCTCCCGAAGACGACGAGTCCCCGAGGACACC 1122

Db 119 GAAGACGAGAAGATGATGATGATGACGACGAAAGAGATGATGATGATGATGAC 178

QY 1123 GAGCAGCATGATCCAACTCGGATCTGA 1151

Db 179 GACGACGATGATGATGACGAAAGATGATGA 207

RESULT 7

LOCUS BG345158

DEFINITION BG345158 865 bp mRNA linear EST 22-OCT-2001

HVSMeg0018018f Hordeum vulgare pre-anthesis spike EST library

HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone

HVSMeg0018018f, mRNA sequence.

ACCESSION BG345158

VERSION BG345158.2 GI:16318857

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae ; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 865)

AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton ,R.D., Close,S.J., Oates,R. and Main,D.

TITLE Development of a genetically and physically anchored EST resource for barley genomics: Morex pre-anthesis spike cDNA library

JOURNAL Unpublished (2001)

COMMENT On Feb 27, 2001 this sequence version replaced gi:13157487.

CONTACT: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 412

Seq primer: AATTAAACCCTCACTAAAGGG

High quality sequence stop: 856.

FEATURES

source

1. 865

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone="HVSMEg0018018f"

/clone_lib="Hordeum vulgare pre-anthesis spike EST library

HVCNDA0008 (white to yellow anther)"

/tissue_type="pre-anthesis spike"

/lab_host="SOLR"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; plants were grown in the greenhouse at the University of California, Riverside (Fenton, SJ Close, TJ Close). Whole spike with awns trimmed were collected at white, green and yellow anther stages (Fenton). Total RNA was prepared from each pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give Bluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close lab (Choi) at the University of California, Riverside.

Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing) Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch , Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html>)"

BASE COUNT 196 a 217 c 271 g 181 t

ORIGIN

Query Match 3.1%; Score 40.2; DB 10; Length 865;

Best Local Similarity 53.5%; Pred.No.3.1;

Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1002 GGACGCAAGAACGAGCAAGAGGCCCTCCAGAAGACTCGAGAGCAGACATGCAGCC 1061

Db 65 GGACGAGACCCGAGGCGCGAGAGCAAGATGAAGACGAGACCAAGACGAGCAGCA 124

QY 1062 AGAGGCTTCTGGAGAAATCTGCGCCCTCCCGAAGACGAGCAAGTCCCCGAGACAC 1121

Db 125 CGATGACGACGATGATGATGATGACGACGACGACGACGACGACGACGACGACCA 184

[illegible]

RESULT	11	558 bp	mRNA	linear	EST 13-JAN-2001
LOCUS	BF825826/c				
DEFINITION	BF825826				
ACCESSION	MR2-HN0036-171100-002-e03	HN0036	Homo sapiens	CDNA, mRNA	sequence.
VERSION	BF825826				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 558)				
	Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.				
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
MEDLINE	20202663				
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HN0036-171100-002-e03&t3=2000-11-17&t4=1) Seq primer: puc 18 forward High quality sequence stop: 254. Location/Qualifiers 1.558 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="HN0036" /dev_stage="Adult" /note="Organ: head_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription o tissue mRNA and cDNA amplification were performed under low stringency conditions."				
FEATURES					
source					
BASE COUNT	96 a	180 c	124 g	158 t	
ORIGIN					
Query Match	3.0%;	Score 39.6;	DB 10;	Length 558;	
Best Local Similarity	45.9%;	Pred. No. 3.8;			
Matches 135;	Conservative 0;	Mismatches 159;	Indels 0;	Gaps 0;	
QY	870	GGGATACGAAGATATTCTGCACGCGCTGGAATATTGCTGAGGAAAGAAATCCTAGCGC	929		
Db	397	GGGACAGCCCAAGAGAGATGGGCTCTGTTGAAGCCCCCTTAAGACAATAAGTCTCCTGAG	338		
QY	930	GCCAGACCCCTGCTCCAGATAGCGTCCCGCAGAAATATCCCGCTGAACCAAGAAAGCCGA	989		
Db	337	TCCAAACCATTTGCTGGAAGCTGACCAAAAGAGAGCTGTCCAGACAGCTGTTAAGAGGG	278		
QY	990	AGGGCGCACCCCGAGCGCAGAAAGCAGCGAAAGAGAGGCCCTCCAGAGACTCGAGAGA	1049		
Db	277	ACCCAGTCCCTAGTAATGCTCTCGAGGCAAGAGCGCCACGACCTGGCAGTGATGA	218		
QY	1050	CGACATGCAGGACAGAGGCTTGTGAGAAATCCTGCGCCCTCCCGAAGACGACGAAGT	1109		

Db 217 GGAAGAGGAGAGAACTCTGAGAAGATGGTATGGTGAACCAACGGGAGCTCTGGGG 158

QY 1110 CCCCAGAGACACCGACGATGATCCAACTCGGATCCTGACTATTACAATGA 1163

Db 157 CTCCGAGGACGATGCTGATACGGTAGATGACTATGAGAGCTGACTCCAACCTCTGA 104

RESULT 12

AV896003/c 567 bp mRNA linear EST 09-NOV-2001

LOCUS

DEFINITION AV896003 Nori Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone rcia43115 3', mRNA sequence.

ACCESSION AV896003

VERSION AV896003.1 GI:16885099

KEYWORDS EST.

SOURCE Ciona intestinalis.

ORGANISM Ciona intestinalis

REFERENCE 1 (bases 1 to 567)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source 1.567

/organism="Ciona intestinalis"

/db_xref="taxon:7719"

/clone="rcia43115"

/clone_1lb="Nori Satoh unpublished cDNA library, young adult"

/tissue_type="whole animal"

/dev_stage="young adult"

BASE COUNT 140 a 148 c 87 g 192 t

ORIGIN

Query Match 3.0%; Score 39.6; DB 9; Length 567;

Best Local Similarity 50.0%; Pred. No. 3.8;

Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 967 CCCGCTGTACCAAGAAAGCGGAGCGCCACCCCGGAGCAGAGAAGCAAGCAAGAAAG 1026

Db 480 CCAACTGAAGAGATCTCAGGAAGAGGAGACAAGAGGAGCAAGGGGAGAGAGGAA 421

QY 1027 GCCCCCTCCAGAACTCGGAGGACGACATGCAGCAGAGAGGCTTCTGGAGAAATCCTGCC 1086

Db 420 GAAGAGGAAGACGATGATGATGACGACGATGACGAAGATGATGATGACGAAGATGAT 361

QY 1087 GCCCTCCCCGAAGACGAGAGTCCCCGAGACACCGAGCAGATGATCCAACTCGGAT 1146

Db 360 GACGAAGATTAATGACGACGATGATGACGAAGATGATGACGACGATGATGACGAAGAT 301

QY 1147 CCTGACTATTACAATGAC 1164

Db 300 GACGACGACGAAGATGAC 283

RESULT 13

AV901789/c 637 bp mRNA linear EST 09-NOV-2001

LOCUS

DEFINITION AV901789 Nori Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone rcia52h17 3', mRNA sequence.

ACCESSION AV901789

VERSION AV901789.1 GI:16890887

KEYWORDS EST.

SOURCE Ciona intestinalis.

ORGANISM Ciona intestinalis

REFERENCE 1 (bases 1 to 637)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source 1.637

/organism="Ciona intestinalis"

/db_xref="taxon:7719"

/clone="rcia52h17"

/clone_1lb="Nori Satoh unpublished cDNA library, young adult"

/tissue_type="whole animal"

/dev_stage="young adult"

BASE COUNT 165 a 161 c 93 g 218 t

ORIGIN

Query Match 3.0%; Score 39.6; DB 9; Length 637;

Best Local Similarity 50.0%; Pred. No. 4;

Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 967 CCCGCTGTACCAAGAAAGCGGAGGCGCACCCCGGAGCAGAGAAGCAAGCAAGAAAG 1026

Db 575 CCAACTGAAGAGATCTCAGGAAGAGGAGGAGGAGACAAGGGGAGAGAGGAA 516

QY 1027 GCCCCCTCCAGAACTCGGAGGACGACATGCAGGCGCAGAGGCTTCTGGAGAAATCCTGCC 1086

Db 515 GAAGAGGAAGACGATGATGATGATGACGACGATGACGAAGATGATGATGATGATGATGAT 456

QY 1087 GCCCTCCCCGAAGACGAGAGTCCCCGAGGACACCGAGCAGATGATCCAACTCGGAT 1146

Db 455 GACGAAGATTAATGACGACGATGATGACGAAGATGATGACGACGATGATGACGAAGAT 396

QY 1147 CCTGACTATTACAATGAC 1164

Db 395 GACGACGACGAAGATGAC 378

RESULT 14

AV897918/c 713 bp mRNA linear EST 09-NOV-2001

LOCUS

DEFINITION AV897918 Nori Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone rcia50120 3', mRNA sequence.

ACCESSION AV897918

VERSION AV897918.1 GI:16887016

KEYWORDS EST.

SOURCE Ciona intestinalis.

ORGANISM Ciona intestinalis

REFERENCE 1 (bases 1 to 713)

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source 1.713

location/Qualifiers

BASE COUNT	ORIGIN
197 a	<pre> /organism="Ciona intestinalis" /db_xref="taxon:7719" /clone="rciad50120" /clone_lib="Nori Satoh unpublished cDNA library, young adult" /tissue_type="whole animal" /dev_stage="young adult" </pre>
174 c	116 g 226 t

Query Match	3.0%;	Score 39.6;	DB 9;	Length 713;
Best Local Similarity	50.0%;	Pred. No. 4.2;		
Matches	99;	Mismatches	99;	Indels 0;
				Gaps 0;

QY	967	CCCGCTGTAA	CCAGAAAGCGG	AGGCGCACCCCGAGCGCAGAAAGCAGCGAAAGAAAG	1026
Db	583	CCA	ACTGAAGAAGAT	ACTCAGGAAGAGGAGGACCAAGAGGGACAAAGGGGAGAAAGAGAA	524
QY	1027	GCCCC	TCAGAAAGACTCGGAGGACGACATGCAGG	CAGAGGCTTCTGAGAAAAATCCTGCC	1086
Db	523	GAA	GACGAAGAAGACGACGATGATGATGACGACGATGACGAAGATGATGATGACGAAGATGAT	464	
QY	1087	GCCCT	CCCCGAAGACGACGACGAAGTCCCCGAGGACACCGACGACGATGATCCAAACTCGGAT	1146	
Db	463	GAC	GAGATTAATGACGACGATGATGACGAAAGATGATGACGACGATGATGACGACGAAAGAT	404	
QY	1147	CCT	GACTATTACAATGAC	1164	
Db	403	GAC	GACGACGACGAAGATGAC	386	

RESULT	15				
AV892667/c					
LOCUS					
DEFINITION					
AV892667	715 bp	mRNA	linear	EST 09-NOV-2001	
AV892667	Nori Satoh unpublished cDNA library, Young adult	Ciona			
intestinalis cDNA clone rc1ad33d12	3'	mRNA sequence.			

ACCESSION	AV892667
VERSION	AV892667.1
	GI:16881763

SOURCE	ORGANISM
intestinalis.	ciona
intestinalis	ciona

ORGANISM
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 715)

AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin-1, T.
TITLE Expressed genes in *Ciona intestinalis*

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh
Department of Zoology

Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp

FEATURES	Location/Qualifiers
source	1. . 715

BASE COUNT	ORIGIN
196 a	/organism="Clona intestinalis"
171 c	/db_xref="taxon:7719"
117 g	/clone="rciad33d12"
230 t	/clone_lib="Nori Satoh unpublished cDNA library, young adult"
	/tissue_type="whole animal"
	/dev_stage="young adult"
	1 others

Query Match	3.0%;	Score 39.6;	DB 9;	Length 715;
Best Local Similarity	50.0%;	Pred. No. 4.3;		
Matches 99; Conservative	0;	Mismatches 99;	Indels 0;	Gaps 0;

QY 967 CCCGCTGTACCAAGAAGCGGAAGGGCGCACCCCGGACGCAGAAAGCAGCGAAAAAG 1026

Db	590	CCAACTGAAGAAGATACTCAGGAAGAGGAGAGACAAGAGGGAGAGAGGAA	531
QY	1027	GGCCCTCCAGAAGACTCGAGGACGACATGCAGGCAGAGCTTCTGGAAAAATCCTGCC	1086
Db	530	GAAAGAGGAAGACGACGATGATGATGACGACGATGACGAAGATGATGATGACGAAGATGAT	471
QY	1087	GGCCCTCCCGAAGACGACGAAGTCCCCGAGGACACCGACGATGATCCAAACTCGAT	1146
Db	470	GACGACATTAATGACGACGATGATGACGAAGATGATGACGACGATGACGACGAAGAT	411
QY	1147	CCTGACTATTACATGAC	1164
Db	410	GACGACGACGAAGATGAC	393

search completed: September 29, 2002, 00:52:16
job time: 10339 sec

THIS PAGE BLANK (USPTO)

OM of: US-09-994-064-11 to: GenEmbl:* out_format : pfs
Date: Sep 29, 2002 11:05 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool/US09994064/runat_27092002_185413_13630/app-query.fasta.1.499
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09994064_ecgnl_1_7551
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-994-064-11
Query length: 434
Database: GenEmbl:*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 2164.800000

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat:AR093367	2311.00	3647.60	8.2e-195	1305	AR093367 Sequence 19 from pate
gb_pat:AR135449	2311.00	3647.60	8.2e-195	1305	AR135449 Sequence 19 from pate
gb_vi:ILU28832	2211.00	3620.76	2.6e-193	18912	U28832 Infectious Iarngotrat
gb_vi:HSMDGLXCO	2226.00	3510.60	3.5e-187	1627	L31965 Gallid herpesvirus 1 gl
gb_pat:A22122	648.00	1000.50	2.3e-47	3502	A22122 ILTV gp60 gene. 6/1994
gb_pat:I04278	199.00	296.37	3.8e-08	1594	I04278 Sequence 2 from patent
gb_pat:I07884	199.00	296.37	3.8e-08	1594	I07884 Sequence 2 from patent
gb_vi:HSIGDB	199.00	296.37	3.8e-08	1594	K02372 Herpes simplex virus ty
gb_pat:E00395	198.00	294.78	4.7e-08	1594	E00395 DNA coding herpes simpl
gb_pat:E00401	198.00	294.78	4.7e-08	1594	E00401 Herpes simplex virus-1
gb_pat:I00103	197.50	293.90	5.2e-08	1608	I00103 Sequence 1 from patent
gb_pat:AR043518	193.50	287.56	1.2e-07	1608	AR043518 Sequence 3 from paten
gb_pat:E00204	193.50	287.56	1.2e-07	1608	E00204 DNA sequence encoding g
gb_pat:I01122	193.50	287.56	1.2e-07	1608	I01122 Sequence 1 from patent
gb_pat:I01496	193.50	287.56	1.2e-07	1608	I01496 Sequence 1 from patent
gb_pat:I07846	193.50	287.56	1.2e-07	1608	I07846 Sequence 1 from patent
gb_pat:I59660	193.50	287.56	1.2e-07	1608	I59660 Sequence 3 from patent
gb_vi:HSIGD	193.50	287.56	1.2e-07	1608	J02217 HSV1 glycoprotein D gen
gb_pat:I70141	190.00	284.92	1.7e-07	1204	I70141 Sequence 1 from patent
gb_vi:HSIDF	190.00	284.92	1.7e-07	1206	L09242 Herpes simplex virus ty
gb_vi:HSIDKOS	190.00	284.82	1.7e-07	1216	L09243 Herpes simplex virus ty
gb_vi:HSIRID2M	190.00	284.77	1.7e-07	1222	L09244 Herpes simplex virus ty
gb_vi:HSIRID1M	190.00	284.56	1.7e-07	1248	L09245 Herpes simplex virus ty
gb_pat:I01099	190.00	283.12	2.1e-07	1440	I01099 Sequence 1 from patent
gb_pat:E02509	190.00	282.99	2.1e-07	1459	E02509 DNA encoding surface gl
gb_pat:E03023	190.00	282.99	2.1e-07	1459	E03023 DNA encoding surface pr
gb_pat:E03111	190.00	282.99	2.1e-07	1459	E03111 DNA encoding surface pr
gb_vi:HEHSV1SU	190.00	261.04	3.5e-06	12979	X02138 Herpes simplex virus t
gb_vi:HSIUS	190.00	253.97	8.7e-06	26245	L00036 Human herpesvirus stra
gb_vi:HE1CG	190.00	236.32	8.4e-05	152261	X54361 H. simplex virus
gb_vi:HSV1ANG	187.00	279.63	3.3e-07	1185	AF333383 Human herpesvirus 1 g
gb_vi:AF333383	186.50	279.52	3.3e-07	1185	AF333383 Human herpesvirus 1 g
gb_vi:HSEGLYPD	184.50	270.01	4.1e-06	2229	M62923 Equine herpesvirus type
gb_vi:HSEE249GD	184.50	258.87	5.5e-07	1185	AF293614 Human herpesvirus 1 p
gb_pat:AR043524	182.50	270.49	1.1e-06	1550	AR043524 Sequence 10 from pat
gb_pat:I59666	182.50	270.49	1.1e-06	1550	I59666 Sequence 10 from pat
gb_vi:HSEGLYD	178.50	265.10	2.1e-06	1409	M59773 Equine herpesvirus glyd
gb_vi:HSEGP1718	178.50	263.81	2.5e-06	1602	M60946 Equine herpes virus gly
gb_pat:E54545	178.50	249.90	1.5e-05	6402	E54545 Herpesvirus recombinant

gb_vi:HSECOMGEN	178.50	218.22	0.0009	150223	M86664 Equine herpesvirus
gb_vi:PVI271966	170.00	250.41	1.4e-05	1590	AJ271966 Pseudorabies virus
gb_vi:PVI271967	170.00	249.01	1.7e-05	1827	AJ271967 Pseudorabies virus
gb_pat:E03567	166.50	246.82	2.2e-05	1308	E03567 DNA encoding fused p
gb_pat:E03115	166.50	238.74	6.2e-05	2925	E03115 DNA encoding fusion

seq_name: gb_pat:AR093367

seq_documentation_block:

LOCUS AR093367 1305 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 19 from patent US 6001369.
ACCESSION AR093367
VERSION AR093367.1 GI:10020117
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1305)
AUTHORS Cochran,M.D. and Junker,D.E.
TITLE Recombinant fowlpox viruses and uses thereof
JOURNAL Patent: US 6001369-A 19 14-DEC-1999;
FEATURES
source location/Qualifiers
1..1305

BASE COUNT

331 a 349 c 344 g 281 t
ORIGIN

alignment_scores:

Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-994-064-11 x AR093367 ..

Align seg 1/1 to: AR093367 from: 1 to: 1305

1	MethisArgProHisLeuAlaArgHisSerArgTyrAlaLysGly	17
1	ATGCACCGTCTCATCTCAGACGGCAGCTCGCTACTACGGAAGAGA	50
17	uValLeuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyA	34
51	GGTGCTTAACAACACATGATGCGGTGGAACGGTGCTGCTCAGCGC	100
34	laAlaValPheThrLeuPheThrPThrCysValArgIleMetArgLHis	50
101	CAGCTGATTACACTCTTTCTGGACTTGTGTACAGATATGCGGAGCAT	150
51	IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl	67
151	ATCTGCTTGTACGCAACGCTATGACCGCCATTTATTTTGAGGAATGC	200
67	apHetrPThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA	84
201	TTTTTGACTATCGTACTGCTTTCTTCCTTCGCTAGCCAGACCGCGC	250
84	laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr	100
251	CCGTCACGTACGACTACATTTTAGCGCGTCCGCCCTCGACCGCTAAC	300
101	IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG	117
301	ATACCGCGGCTTGCGCCGTATACAGATACCTACAGGATATCAAGAGG	350
117	YCysAspValValGlyLeuAsnProIleSerAsnValAspAspMetIleS	134
351	CTGCGACGTTGTGAGCTCAACCCGATTTCTAACGTGACGACATGATAT	400
134	erAlaAlaLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyL	150
401	CGCGCGCCAAAGAAAGAGAAGGGGGCCCTTCGAGGCCCTCCGTCGTC	450

151 trpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProIl 167
|||||
451 TGGTTCACGTGATTAAAGGGCGACGACGGCGAGGACAAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
|||||
501 CTATAGAAAGAGTACAGGGAATGTGGCGACGTACAACTGCTATCTGAAT 550
184 ysaIaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
|||||
551 GCCCGCTCAATCTGCACAGATGTGGCAGTGGACTATGTCTCCTAGCAC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
|||||
601 CTGTATCGCGAAATGGCGCGGACTGACTATATTCTCCCCCAGTGTGC 650
217 aleuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
|||||
651 GCTCTGTGGCCAATACTTGTGACCTGAATAATCGGAGATTGTGGCCAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
701 CAGCTCTGTACTTCTAGAAAGTTAACGATCGCTGTTAAAGATCGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
|||||
751 CAGCTTAACCTTTTACCGTCGAATATGCTGACCAACAAGACAGATATCAG 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
|||||
801 TGGATTTCAGGCGAACAACCTTATCCGATCGCAGACACCAATACAGCAC 850
284 tsAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTyrAsn 300
|||||
851 ACGCGGACGACGTATATCGGGGATACGAAGATATTCTGCAGCGCTGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
|||||
901 AATTTGCTGAGAAAAAGAAATCTTAGCGCGCCAGAACCTCTGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
|||||
951 CGTCCCGCAAGAAATCCCGCTGTAAACCAAGAAAGCGGAAGGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
|||||
1001 CGGACGCGAGAAAGCAGAAAAAGAGCCCTCCAGAAAGACTCGAGAGAC 1050
351 AspMetGlnAlaGluAsnSerGlyGluAsnProAlaAlaLeuProGluAs 367
|||||
1051 GACATGCAGGACAGAGGCTTCTGAGAAAAATCCTGCCGCCCTCCCGAAGA 1100
367 paspGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
1101 CGACGAAGTCCCCGAGGACACCGACGACGATGATCCAAACTCGGATCCTG 1150
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
|||||
1151 ACTATTACAATGACATGCCCCCGCGGTGATCCCCGGTGAGAGAGACTACTAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaIaIaPheValaIaCysAl 417
|||||
1201 AGTCTAATGCCGTCTTCATGCCCATATATTCGCGCGCTTCGTAGCCTGCC 1250
417 aValAlaLeuValGlyLeuLeuValTyrSerIleValLysCysAlaArgS 434
|||||
1251 GGTGCGCGCTGTGGGCTACTGTTTGAGAGCATGTAATAATGCGCGCGTA 1300
434 er 434
||
1301 GC 1302

seq_name: gb_pat:AR135449
seq_documentation_block:
LOCUS AR135449 1305 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 19 from patent US 6136318.
ACCESSION AR135449
VERSION AR135449.1 GI:14476121
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1305)
AUTHORS Cochran, M.D. and Junker, D.E.
TITLE Recombinant fowlpox viruses and uses thereof
JOURNAL Patent: US 6136318-A 19 24-Oct-2000;
FEATURES
source location/Qualifiers
BASE COUNT 331 a 349 c 344 g 281 t
ORIGIN
alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-994-064-11 x AR135449 ..
Align seg 1/1 to: AR135449 from: 1 to: 1305
1 MethIsArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyI 17
|||||
1 ATGCACCGCTCCCTCATCTCAGACGGCACTCGCGTTACTACGCGAAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyLysArgCysSerGlyA 34
|||||
51 GGTGCTTAACAACAACATGATGATGGCGTGGAAGACGGTGTGCTCAGGCG 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
|||||
101 CAGCTGATTCACCTCTTCTGTGACTTGTGTACAGATTATGCCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
151 ATCTGCTTTGTACGCAACGCTATGACCGCCATTATTTTTGAGGAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
|||||
201 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGAGACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
251 CCGTCACGTACGACTACATTTTAGGCCGTGCGCGCTGAGCGGCTTAACC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
|||||
301 ATACCGGCGGTGGCCCGTATAACAGATACTCAGGATATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
|||||
351 CTGCGACGTTGTGCAAGCTCAACCCGATTTCTAACGTGAGCAGACATGAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyPropheGluAlaSerValVal 150
|||||
401 CGGCGGCCAAAGAAAAAGAGAGGGGGGCCCTTTCGAGGCCCTCCGTCGTC 450
151 trpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProIl 167
|||||
451 TGGTTCTACGTGATTAAGGGCGAGCGAGCGGAGGACAAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184

|||||
501 CTATAGAAAAAGATACAGAGGAAATGTGGCGACGATACAACTGCTATCTGAAT 550
184 ysalavalinsertalaglnmetrrpalavalasptyrvalproserthr 200
551 GCGCGCTTCAATCTGCACAGATGTGGCGAGTGAGCTATGTTCTTAGCACC 600
201 leuvalserargasnlglyaleuThrlepheserProThrAla1 217
601 CTTGTATCGGAATGGCGGGGAGTACTATATTTCTCCCACTGCTGC 650
217 aleuSerglyIntyrleuThrleuLysileglyargpheAlaInt 234
651 GCTCTGTGGCAATACTGTGACCCCTGAAAATCGGGAGATTGGCCAAA 700
234 hralaleuvalThrleuGluValasnasparGcysleuLysileglyser 250
701 CAGCTCTCGTAACCTAGAAATTAAAGATCGCTGTTAAAGATCGGGTCG 750
251 GlnleuAsnpheluProserLysCysTrpThrThrgluGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCGTCGAAATGCTGGACAACACAGATATCAGAC 800
267 rglYpheGlnGlyLuhIsleuTyrProIleAlaasprhrasnThrArgH 284
801 TGGATTTCAAGCGGAACACTTATCCGATCGCAGACACCAATACACGAC 850
284 tsAlaaspaspvalTyrargGlyTyrGluaspIleleuGlnArgTrpAsn 300
851 ACGCGGACGACGTATATCGGGGATACGAAGATATCTGCAGCGCTGGAAT 900
301 AsnleuLeuArgLysLysAsnProSeralProaspProargProaspse 317
901 AATTGCTGAGAAAAAGATCTAGCGCGCCAGACCCTCGTCCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrp 334
951 CGTCCCGCAAGAAATTCCTGTAACCAAGAAAGCGGAGCGCCACACC 1000
334 roAspAlaGluSerSergLysLysAlaProProGluaspSergLyuasp 350
1001 CGGACGCGAAGACGCGAAAGAGGCCCCCTCCAGAGACTCGGAGGAC 1050
351 AspmetGlnAlaGlnAsnserGlyLyuasnProAlaAlaLeuProGluas 367
1051 GACATGCAGCGAGGCTTGTGAGAAAAATCCTGCCGCCCTCCCGAAGA 1100
367 paspGluValProGluaspThrGluHisaspaspProasnSeraspProa 384
1101 CGAGCAAGTCCCCGAGGACACCGAGCAGATGATCCAACCTCGGATCCTG 1150
384 sPTyrrTyrAsnaspmetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGCCCGCGGTGATCCCGGTGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSermetProIlepheAlaAlapheValAlaCysAl 417
1201 AGTTCTAATGCCGTCTCCATGCCCATATTCCGCGCGCTCGTAGCCGCGC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGCGCTCGTGGGCTACTGTTGGAGCATCGTAAATGCGCGCGTA 1300
434 er 434
1301 GC 1302

seq_name: gb_v1:ILU28832

seq_documentation_block: 18912 bp DNA linear VRL 09-AUG-1996
LOCUS ILU28832
DEFINITION Infectious laryngotracheitis virus US10, US2, protein kinase, UL47,
glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein

ACCESSION E, ORF9 genes, complete cds.
VERSION U28832
KEYWORDS U28832.1 GI:1486484
SOURCE Infectious laryngotracheitis virus.
ORGANISM Gallid herpesvirus 1
REFERENCE 1. (bases 1 to 18912)
AUTHORS Wild,M.A., Cook,S. and Cochran,M.
TITLE A genomic map of infectious laryngotracheitis virus and the
sequence and organization of genes present in the unique short and
flanking regions
JOURNAL Virus Genes 12 (2), 107-116 (1996)
MEDLINE 97033380
REFERENCE 2 (bases 1 to 18912)
AUTHORS Wild,M.A.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1995) Martha A. Wild, Syntro Research Laboratory,
3535 General Atomics Court, San Diego, CA 92121, USA
FEATURES
source location/Qualifiers
1. 18912
/organism="Gallid herpesvirus 1"
/strain="USDA challenge strain"
/db_xref="taxon:10386"
<1. 2909
/note="large inverted repeat flanking the short unique
region"
/rpt_type="inverted
697. 1533
/codon_start=1
/product="US10"
/protein_id="AAC55093.1"
/db_xref="GI:1486485"
/translation="MENMLDGCYPLALMDSDHITAHAVPGRGERRROGAAVASSESADS
VDPICIRIASRLMRELVLEISSELDKDYGEFTSARDRNALIAANERLSAFLGASRATR
GIGLRPRWASTESVANSPTDPNNGLGELEAMEGIEGDFWLDSDGREDESRTM
QSENMRFVIEKELLSWLSRHLPADLASARETSSRLAAGHWCCLMHPRPCREACLYD
SIYVQSLFCVGTGRVPOSEMRREYLAALRAGAAANSPEVSASIFARDAGIALALAR
RR"
complement(1900. 2784)
/note="coding region unique to avian herpesviruses"
/codon_start=1
/product="unknown"
/protein_id="AAC55094.1"
/db_xref="GI:1486486"
/translation="MSKCYCLARHLKSPRCVGRVAFGGLATMSRPPTSHLDLAFSA
AFRGTDLPGRFWRASQSCDIFFWPDLAIVQARAYFEGERLGSQVAEDITAND
PRIAPAKRVAVAAGVLTALSELVGELEESKVGKQIPRAAMEIRDPKVP
GPDILSFSAAVELPVLIRARGGAHSRAHWNQSSAPAPAGLAIRIGMEMVRSLL
IALPLSNFTLEBDLPEGSQNSIRAFVAHLMNCVATDKIMSPDVRVVEESFYSHCLRE
IIMCERAFCYPCNPPEPKW"
2910. 16003
/note="short unique region"
complement(2916. 3605)
/codon_start=1
/product="US2"
/protein_id="AAC55095.1"
/db_xref="GI:1486487"
/translation="MAPVKVTIVSAVDSHYKLPNSRFEI SDSGKELVHAKTMASYD
RPSTLSVIVRPASLYEVSGELFSLPRMCRVIRFEGEGDPGVSPEWSGLDAGFYHLS
SGAYAAKEFHLWVLTADICMAALNLPAPKTELLITETGKNFERGVEIFLVNGDKTTL
SLSHPSVWTTLPASSLRTWPYSTVYKPKVPNSAAYCVSDSDSDBERQPKFFLGSLEK
SKKPRSPRRR"
3694. 5124
/codon_start=1
/product="protein kinase"
/protein_id="AAC55096.1"
/db_xref="GI:1486488"
/translation="MRFRRICSRRAEKRRRTTENPLTSKRVCYUDSFSRTMSLRPYA
EILPTAEGVERLAELVSVTMTERAEPYENTAVNSIPANENGONFAYAGDSPSTTEK
VDGSHTDEEASSDYAGVPVLAQTRKHSDEFIQHFRVLDDLVEGAYGFCIDVRRYTE
EEQRRRGVNSTNGKSKCKRLIAKYVKNGTAAASQLENEILVLGRLNHENVLKIQEIL

CDS
/codon_start=-1
/product="U147"
/protein_id="AAC55097.1"
/db_xref="GI:1486489"
/translation="MTLPHRLTKRPFARFRCVFIHYSETKLDRYKNTMLLYRPDST
MRHSGDANHRGIRPRRKSIGAFSAREKTGRNALTESSSSSDMLDPFSTDFEGGKW
TVDGPADITAEVLSQAMDVLQVVKHEDAEERVIYESKPTPIQPFNAMPDGSPSWAOD
FTRAPIVYPSAEVLDALVKGAFVSRLQCVPTRSKKSVTVRDAQSFLGDSFWRIM
QNVYTVLRQHTRLRHPSSKSIVNCNDPLWYANQPHWRGMVPSLKASPEENI
QHGPMAAVERNAGAGFLMPAMRAAEERDKRLRACLSSLDIMDAAVLASPEFYWRG
VQDTSREFPALGCLSEYFALVLAETVLATMEPHALVFMRALADGNEDYDetryid
PVKNEYLNGAEGTLRGIVASNTALAVVCANTYSTIRKLPSVATSACNVAYRTEYLKA
RRPMSDIYRILQKEFFFYIAWLQRVATHANECNLILKRSVDTGAPPLFRASSEKRL
QOLNMLCPLVPIQYEDFESKAMGSELKREKLETFYKAISSDRDPGSLRFLISDHAR
EIIADGVRFKPVIDEVRAVSVALSTAAAGVKARLLTSVRAPVPGAGAVSARRKSEI"
7245. .8123
/codon_start=1
/product="glycoprotein G"
/protein_id="AAC55098.1"
/db_xref="GI:1486490"
/translation="MSGFSNIGSIATVSLVCSLCLGASVIGAPVLDGLESPFPFGKI
IAQACNRTTIEVTPWMSDYSGRTEGVSVEVKWFYGNMPESEFVGDSSETSGHEDLS
TCWALIHNLINASVCRASDAGIPDEFQCEKVQRLRSGVELGSYVSGNSLVLYPGMY
DAGIYAYQLSVGKGYSYILVDPNPGCHDOYGYTYYSIADASDLSYDVASPEL
DGPMEDYSNCLDMPPLRPMTVCSHDVEEOENATDELYLWDEECAGPLIDEYDERSE
TMPRMVVFSPSTLQQ"
8333. .11290
/note="ORF5"
/codon_start=1
/protein_id="AAC55099.1"
/db_xref="GI:1486491"
/translation="MGTMLVLRLEFLAVADAALPTGRFCRWKVPFGGTIQENLAVLA
ESPVTGHATYPPREGAVSFQIFADPTLRIRYGATEDELALEGTASADADNVTFSL
YRPREIHGAYFTIGVFATGQSTESSYSIVSRVLNASLERSVRLTPCDENLQNEP
TWGSRWLGPSPSYVRNDVAVLTKAQYIGECYSNSAQOTGLSLMTEFYSPKRIYN
VTWTTGGPSPSRITVYSSRENGOVLARNVSDGLVKYTPIDGRAMINVIANSPADS
GSVLAFTAFREGKLPSAIOHLRIMSGTEPGTETTFDCQMIETPYRALGSNVRD
SIRPGATLPPEDTAAPDEFDTPTVPEPAITTLIPRSTSDMGFESTARATGSET
LSVVOETDRTLSTPLTLPTGPESENTLPPTAPGISTEPMSAAHETTOTOSAETV
VFQSPSTESSTARSOSEPWFTQPTSTEQALITQIAETEALETQTPSAEQMTEY
QTPGATEBAPAQTPSTIPEIFTOGSRSTPETARAPSAPEVTOSSSTVTEVETQPS
TVPKTLLSSSTEPAIFTRTQSPAGTAFQTSSAEPDTRQSTETHEFTQAPSTVPKA
TQTPSTEPEVLQSPSTEPVPTRLGAPEITQTPSAPEVYTRSSSTMPETAOSTP
LASONPTSSGCTGHTNTEPRTPVQTPHTQKLYTENKTLSEPTVVSSEHEMSTAESOT
PLLDVAKIVEKFSNDGEVATCVSYKSPYRVETNMKVDLVDMDEISGNSPAGVENS
NEKWKOLYRVTGDRTSVQIMCLSCTSHPPEYCLFDTSLAREKDIAPELYFTSDP
QYAYCTILPSGVVPRFEWSLNNVSLPEYLITATTVSHTAGOSTVYWKSSARAGEAWIS
GRGNIVECTVLISDGTRVYTRKERCLTNWIAVENGAQAQLYSLFSGLVSLGCSI
SALYATLWTAIYF"
11098. .12402
/codon_start=1
/product="glycoprotein D"
/protein_id="AAC55100.1"
/db_xref="GI:1486492"
/translation="MHRPHLRHRHRYAKGEVLNKHMDCCGKRCCSGAAVFTLFWTCV
RIMREHICFYRNAMDRHLRLRNHFVYLSSFASOSTAAVYDYILGRALDALITP
AVGPYNRYLTRYSRGCDVVELNPISNVDDMISAKEKEKGPEEASVWVYVIKGDG
EDKYCPIYRKEYREGCDVQLSECAVQSAQMAVDYVPSLVSRNAGLTIESTPAAL
SGOYLTLTKIGRFAOTALVTLVENDRCLKIGSQLNFLPSKCWTEYOQTOGGEHLYP
IADTNRHADVYRGYEDILQRMNNLIRKKNPASDPDRPDSVQEIPTAVTKKAEGRTF
DAESSEKKAPPESEDDMOAASGENPALPEDDEVEDTEHDDPSDDPDYNDMPAV
IPVEETTKSSNAVSMPIFAAFVACAVALVGLLVMSIVKCARs"
12510. .13598
/codon_start=1
/product="glycoprotein I"
/protein_id="AAC55101.1"
/db_xref="GI:1486493"

/translation="MASLGTLLALLAATLAPFGAMGIVITGNHVSARIDDDHIYIAP
RPEATIQQLFEMPQGRPHKPYSGTVRAVAFKRSDDITNOCYOELSEERFENCTHRSSVF
VGKATYEYTFASANRLTGPPHPEKLTIRNRPNDSGMYIVRLDDTKEPIDVPAIQ
SVYQFANTAAATRGLYSKASACRYTGLPTVOLRAYLRTESWRNWOAYVATEATTSAEA
TTPTPYATASASELAEHFTFPWLENGVDHYEPTPANENSNVTVRLGTMSPTLIGTV
AAVVSATIGLVIVISIVTRNMCPTPHRLKLDTVSQDDEERSQTRRESRKFGPVACELNK
GADQDSELVELVAIVNPSALSSPDSIKM"
13792. .15291
/codon_start=1
/product="glycoprotein E"
/protein_id="AAC55102.1"
/db_xref="GI:1486494"
/translation="MNLVTVLACGLARLTATRHVLFLEGTQAVLGEDDDPRNPEGT
VIKWTIVLRNACKMKAADVCCSPNYCFHDLIYDGKKDCPPAGPLSANLVILLKRGES
FVVLGSLHNSNITNIMTEYGLLFDVPTRSDEGIYFRRTSQPDLAMETTSYNVSVL
SHDEKAPAPHEVEIDTIKPSFAHAEHLQMLPFHELNDNSPTVTVPLRVFPTEHV
KFNVTYSWYGFEDVKEECEEVKLFEPVYHPTDGKQCFPATNQCLIGSVLAEFLGA
SLDCSRDLTEDCHENRVNPLRFSRSESAGLVISPLIATPKVLIIVSDGDLIGW
SYTVLGRNSBRVVEETHMPSKVPMKVVIGSPGMDETGNKMYEVAVGAATCVIL
TCALLVGKKCPAHQMGTFSKTEPLVAPLPKNEFEAGGLTDSEVIYDEVYEPFLFRG
CKOEFREDVNTFFGAVVEGERALNFKSAIASMADRILANKSGRRNMSDY"
15298. .16080
/note="ORF9"
/codon_start=1
/protein_id="AAC55103.1"
/db_xref="GI:1486495"
/translation="MPFKTRGAEDAAAGKRFKKSRRNEILPTRLNGTKTAGLSN
TOPIPWNPKEFSARGESDNHACKDTEFYRTCCASRSTVSSQPDSPHTMPTEYGRVS
AKRKKLSSSDCGAHPQVSKLPDSQAAPARTYSSAORYTVDEVSPTPGDAVAD
LETRAELPGATTEQOTESKNKLPNQOSRLKPKPTNEHVGERCPSESTEAPSLGILSR
VGAIIANELARMRACCLPLAASAAAAGIYAAARAALOKQGR"
16004. .>18912
/note="large inverted repeat flanking the short unique
region"
/rpt_type=inverted
16129. .17013
/note="coding region unique to avian herpesviruses"
/codon_start=1
/product="unknown"
/protein_id="AAC55104.1"
/db_xref="GI:1486496"
/translation="MSKCYCLARHLKSPRCVGRVAFEGGLATMSRPTSHLDLAFSA
AFRGTDLPGRFWRASOSCDIFFWPDLAAYVQAARAYFEGKERLGSLOVAEDITAD
PRIAPAKRAVAAVGLWTALSELVGGPNGELESKVGKQIPRAAWETRDVPKVPVI
GPDILSFSAAVELPVLYIRARGAHSRAHNNQSSAPAGLAAIRIGMEMVRSLLV
IALPLSNFTLPEDLPEGSONSIRAFVAHLMNCVATDKIMSPDVVRVPESEFSYSHCLRE
IIMCERAFQPCNPPEPKW"
complement(17380. .18216)
/codon_start=1
/product="us10"
/protein_id="AAC55105.1"
alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-994-064-11 x ILU28832 ..
Align seg 1/1 to: ILU28832 from: 1 to: 18912
1 MethisArgProHisIeuArgArghisSerArgTYRtYRAlaIalsglyG1 17
|||||
11098 ATGCACCGTCTCATCTCAGACGCACTCGGCTTACTACGCGAAAGAGA 11147
17 uValIeuasnlYshisMetaspcysgLyglysArgCysCysSerGlyA 34
|||||
11148 GGTGCTTAACAAACACATGATTCGCGTGGAACGGTGCTGCTCAGGCG 11197
34 IaAlaValpherThrIeuphetrPhrcysValArgIleMetArgGluHis 50
|||||
11198 CAGCTGTATTCACTCTTTCTGGACTTGTGTACAGATTATGCGGAGACAT 11247

51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
11248 ATCTGCTTTGTACGCAACGCTATGACCGCCATTATTTTGTAGGAATGC 11297
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThraAla 84
11298 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGACCGCCG 11347
84 IValIThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
11348 CCGTCACGTAACGACTACATTTTAGCCCGTCGCGCGCTCGACGCGCTAAC 11397
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
11398 ATACCGGCGGTGGCCCGTATACAGATACTCCTACTAGGGTATCAAGAG 11447
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
11448 CTGGCAGCTGTGTCGAGCTCAACCCGATTCTAACGTGGACGACATGATAT 11497
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
11498 CGCGCGCCAAAGAAAAGAGAGGGGGCCCTTCGAGGCCCTCCGTCGTC 11547
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
11548 TGGTTCTACGTGATTAAAGGCGCAGCGCGAGAGACAAGTACTGTCCAAT 11597
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
11598 CTATAGAAAAGAGTACAGGGAATGTGGCGACGTACACTGCTATCTGAAT 11647
184 ySalAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
11648 GCGCGCTCAATCTGCACACAGATGTGGCAGTGGACTATGTCTAGCAC 11697
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
11698 CTGTATCGCAAAATGGCGGGACTGACTATATTCCTCCCACTGCTGC 11747
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
11748 GCTCTCTGGCCAATACTGTGTGACCCGTGAAAATCGGAGATTGGCGAAA 11797
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
11798 CAGCTCTCGTAACCTAGAGTTAACGATCGCTGTTAAAGATCGGGTCG 11847
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
11848 CAGCTTAACCTTTTACCCTCGAATCTGTGACACAGACAGTATCAGAC 11897
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
11898 TGGATTTCGAAGCGCAACCTTTATCCGATCGCAGACCAATACACGAC 11947
284 IsAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
11948 ACGCGGACGACGTATATCGGGGATACGAAGATATCTCGACGCGCTGAAT 11997
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
11998 AATTGCTGAGGAAAAAGATCTTAGCGCGCCAGACCCCTCGTCCAGATAG 12047
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
12048 CGTCCCGCAAGAAATTCCTGCTTAACCAAGAAAGCGGAGGCGCACCC 12097
334 roAspAlaGluSerSerGlyLysLysAlaProProGluAspSerGluAsp 350
12098 CGGACGCGAGAAAGCAGCGAAAGAGGCCCTCCAGAGACTCGGAGGAC 12147

351 AspMetGlnAlaGluAlaSerGlyLysProAlaAlaLeuProGluAs 367
12148 GACATGACAGCAGAGGCTCTGAGAAAATCCTGCCGCCCTCCCGAAGA 12197
367 pAspGluValProGluAspThrGlnHisAspAspProAsnSerAspProA 384
12198 CGACGAGATGCCCCGAGGACACCGAGCAGATGATCCAAACTCGATCTG 12247
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrLys 400
12248 ACTATTACATGACATGCCCGCGGTGATCCCGGTGGAGAGACTACTAAA 12297
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
12298 AGTTCTAATGCCGCTCTCCATGCCCCATATTGCGCGCGTTGTAAGCTG 12347
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArg 434
12348 GGTCCGCGCTCGTGGGGCTACTGTTGGAGCATCGTAAATGCGCGGTA 12397
434 er 434
12398 GC 12399
seq_name: gb_v1:HSMDGLYCO
seq_documentation_block: 1627 bp DNA linear VRL 06-MAY-1996
LOCUS HSMGDGLYCO
DEFINITION Gallid herpesvirus 1 glycoprotein D (gpD) gene, complete cds.
ACCESSION L31965
VERSION L31965.1 GI:493595
KEYWORDS glycoprotein D.
SOURCE Gallid herpesvirus 1.
ORGANISM Gallid herpesvirus 1
VIRUSES: dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
REFERENCE 1 (bases 1 to 1627)
AUTHORS Johnson,M.A., Tyack,S.G., Prideaux,C.T., Kongsuwan,K. and Sheppard,M.
TITLE Sequence characteristics of a gene in infectious laryngotracheitis virus homologous to glycoprotein D of herpes simplex virus
JOURNAL DNA Seq. 5 (3), 191-194 (1995)
MEDLINE 95337426
FEATURES
source location/Qualifiers
1..1627
/organism="Gallid herpesvirus 1"
/strain="SA-2"
/db_xref="taxon:10386"
/note="also known as Infectious laryngotracheitis virus (ILT)"
TATA_signal
144..151
/gene="gp"
gene 144..1517
/gene="gp"
326..1441
/gene="gp"
CDS
/codon_start=1
/product="glycoprotein D"
/protein_id="AA98925.1"
/db_xref="GI:493596"
/translation="MDRHLFLRNAFWTIVLLSSFASQSTAAVTVYILGRALDLTI
PAVGPNRYLTRVSRGCELVENPISVNDMISAKKEKGGPFESVSVFYIKGDD
GEDKYCPYRKEYRECGDVOLSECAVSAQMALDYVSSITVSRNGAGLTIFSPNAA
LSGQYLLTLKIGRFAQTALVTLVENDRCLKIGSRNLFPSKCWTEOYOTGFOGHELY
PIADTNTRHADDVYRREYEDILQWNNILRKKNPSADPRPDSVPEIIPAVTKKAEGRT
PDAESSEKRAPDESDDMOAEASGENPALPEDDEVPEDETHDDPNSDPDYNDMPA
VIVPEETKSSNAVSMPIFAAFVACAVASWGWYWGAS"
polyA_signal
1507..1517
/gene="gp"
BASE COUNT 410 a 428 c 428 g 361 t
ORIGIN

alignment_scores: Quality: 2226.00 Length: 434
 Ratio: 5.189 Gaps: 0
Percent Similarity: 98.848 Percent Identity: 97.465

alignment_block:
US-09-994-064-11 x HSMGDLYCO ..

Align seg 1/1 to: HSMGDLYCO from: 1 to: 1627

```
1 MethisArgProHisLeuArgArgHisSerArgTyrAlaLysGly1 17
|||||
155 ATGCACCGTCCATCTCAGACGGCAGCTCGCTTACTACGCGAAAGGAGA 204
17 uValLeuAsnLysHisMetAspCysGlyLysArgCysSerGlyA 34
|||||
205 GGTGCTTAACGAACACATGATTCGGGTGAAACGGTCTGCTCAGGCG 254
34 laAlaValPheThrLeuPheThrPThrCysValArgIleMetArgLuhis 50
||
255 CACGTGTATTCACCTTTCTGAGACTTGTGTACAGATTATGCGGAGCAT 304
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
305 ATCTGCTTGTACGCAACGCTATGAGCCGCCATTTATTTTGAGAGATGC 354
67 aPheTyrThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
|||||
355 TTTTGGACTATCGTACTGCTTCTTCTCCTTCGCTAGCCAGACACGCCGCG 404
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
405 CCGTCACGTACGACTACATTTTAGGCCGTGCGCGCTCGACCGCTAACCC 454
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG1 117
|||||
455 ATACCGGCGGTGGCCCGTATTAACAGATACCTCAGTGGGTATCAAGAGG 504
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIles 134
|||||
505 CTGCGAGCTTGTGAGCTCAACCCGATTTCTAACGCTGAGACATGATAT 554
134 eraAlaAlaLysGluLysGluLysGlyLysProPheGluAlaSerValVal 150
|||||
555 CGCGCGCCAAAGAAAAAGAGAGGGGGCCCTTTCGAGGCCCTCCGTCGTC 604
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProI1 167
|||||
605 TCCTTCTACGTGATTAAGGGGAGACGAGCGGAGACAACTACTGTCCAAT 654
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
|||||
655 CTATAGAAAAAGAGTACAGGGAATGTGGCGAGCTACAACTGCTATCTGAAT 704
184 ySaIaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
|||||
705 GCGCCGTTCAATCTGCACAGATGTGGCATTTGACTATGTCTTCTAGCACC 754
201 leuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
|||||
755 CTGTATTCGCGAAATGGCGCGAGCTGACTATATTCGCCCAATGCTGC 804
217 aleuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
|||||
805 GCTCTCTGGCCATACTGCTGACCCCTGAAATCGGAGATTGCGCAAA 854
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
855 CAGCTCTCGTAATCTAGAAAGTTAACGATCGCTGTTAAAGATCGGGTCG 904
251 GlnLeuAsnPheLeuProSerLysCysTyrThrThrGluGlnTyrGlnTh 267
:::|||||
905 CGGCTTAACCTTTTACCCTGCAAAATGCTGACAAACAGACATATCAGAC 954
```

```
267 rGlyPheGlnGlyLuhHisLeuTyrProIleAlaAspThrAsnThrArgH 284
|||||
955 TGGATTTCAAAGCGCAACACCTTTATCCGATCCGACAGACACCAATACAGCAG 1004
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTyrPsn 300
|||||
1005 ACGCGAGCAGCTATATCGAGATACGAAGATATTTCTGACAGCGCTGGAAT 1054
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
|||||
1055 AATTTGCTGAGGAAAGAAATCCTAGCGCGCAGACCCCTGCTCAGATAG 1104
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
|||||
1105 CGTCCCGCAAGAAATTCGCCGCTGTAAACCAAGAAACGGAGGCGCAGCC 1154
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
|||||
1155 CGGACGCAAGAAAGCAAGTGAAGAAAGAGCCCTCCAGAAAGACTCGAGAGAC 1204
351 AspMetGlnAlaGluAlaSerGlyLysLysAsnProAlaAlaLeuProGluAs 367
|||||
1205 GACATGCAAGCAGCAGAGCTTCTGAGAAATCTCTGCCGCCCTCCCGAAGA 1254
367 pasPgluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
1255 CGACGAAGTCCCGCAGGACACCGAGCAGCATGATCCAAACTCGGATCCTG 1304
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
|||||
1305 ACTATTACATGACATGACATGCCCGCGTGTATCCCGTGGAGGAGACTACTAAA 1354
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
|||||
1355 AGTTCTAATGCCGTCTCCATGCCCATATTCGCGGCGTTCGTAGCCTGCGC 1404
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
|||||
1405 GTGCGCGTGC.GTGGGGCTACTGTTTGAGAGCATCGTAATAATGCGCGCTA 1453
434 er 434
||
1454 GT 1455
```

```
seq_name: gb_pat:A22122
seq_documentation_block:
LOCUS A22122 3502 bp DNA linear PAT 22-JUN-1992
DEFINITION ILTV gp60 gene.
ACCESSION A22122
VERSION A22122.1 GI:583833
KEYWORDS
SOURCE
ORGANISM Gallid herpesvirus 1.
REFERENCE 1. (bases 1 to 3502)
AUTHORS INFECTIOUS LARYNGOTRACHEITIS VIRUS VACCINE
TITLE Patent: WO 9203554-A 51 05-MAR-1992;
JOURNAL Location/Qualifiers
FEATURES
source 1..3502
/gene="ILTV gp60 gene"
/gene="ILTV gp60 gene"
/gene="ILTV gp60 gene"
/codon_start=1
/protein_id="CAA01588.1"
/db_xref="GI:583834"
/translation="MGTMVLVHLFLAVADALPTGRFCRWKVPVPGGTIQENLAVLA
```


ESPVTGHATYPPREGAVSFQIFADTPTLRIRYGADEDELALERGTSASADNVTSLS
YRPRPEIHGAYFTIGVEATGOSTESSYSIVSRVLVNASLERSVRLTETPCDENFLQNER
TRGSKRWLGPPSPYRVNDVAVLTKAQYIGECYSNSAAQTGLTSLMTFFYSPKRIVN
VTWTTGGPSPSRITVYSSRNGQPVLRNVSDBGFLVKYTPDIDGRAMINVIANSPADS
GSVLAFTAFREGKPSVIQLHRIDISGTEPGTEETFDCKMIETPYRALGSNVRDD
SIGPGATLPPEFDAADEFDTGTSPTPTTVPPEAITLLIPRSTSDMGFFSTARATGSET
LSVPQETDRITSTPLTLPLRGESENTLPPAABGISTETPSAHEHTQOSAEIV
VFTQSPSTESSTARASQSEPMWYFTQSTEQALTLQIAETEVLEFTQPSDEQMTFT
OTPGAETEARQPPSTIPEIFETQSRSTPREPARAPSAAPEVFTQSSSTVEVFTQPS
TVPKTTLSSSTEPALFTRTQASGTEAFQTSSAGPDMRTOSTETHTFTQAPSTVPA
TOTPSTEPEVLQSPSTEPVPFIRTLGAPEITQTPSTVPEITQTPSAPEVYTRSSS
TMPETAOSTPLASONPTSSSGTGHNTPEPTVPTQTPHTOKLYTENKTLSPVVSSEF
HEMSTAESQTPLDKIVEVKFSNDGEVATCVSVKSPYRVETNWKVDLVDVMEIS
GNSPAGVFNSEKQOLYRVTDGRTSVQMLCSCTCHSPEPYCLFDTSLAREKDI
APELYFTSDPQTAICTITLPSGVVPREFWSEFNVSLEPYLTATTVVSHTAGSTVWKS
SARAGEAWISGRGNIYECTVLISDSTRVTRKRERCLNTWIAVENGAQAQAHVYSLSF
GLVSGLCGSISALYATLWTAIYF"

BASE COUNT 845 a 969 c 920 g 768 t

ORIGIN

alignment_scores: Quality: 648.00 Length: 123
 Ratio: 5.311 Gaps: 0
Percent Similarity: 99.187 Percent Identity: 96.748

alignment_block:
US-09-994-064-11 x A22122 ..

Align seg 1/1 to: A22122 from: 1 to: 3502

1 MethisArgProHisLeuArgHisSerArgTyrTyrAlaLysGlyL 17
|||||
3134 ATGCACCGTCTCATCTCAGACGGCAGCTCGCGTTACTACGCCGAAGAGAGA 3183
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyA 34
|||||
3184 GGTGCTTAACGACACATGATGCGGTGGAAGGCGGTGCTCAGGCG 3233
34 laAlaValPheThrLeuPheThrCysValArgIleMetArgGluHis 50
|||||
3234 CACGTGTATTCACTCTTTCTGGACTGTGTCAAGATTATGCGGAGCAT 3283
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
3284 ATCTGCTTTGTACGCAACGCTATGACCGCATTTATTTTGAAGAATGC 3333
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
|||||
3334 TTTTGGACTATCGTACTGCTTCTCTTCCTTCGCTAGCCAGACACCGCCG 3383
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
3384 CCGTCACGTTACGACTACATTTTAGCGCGCTCGCGCTCGACCGCTAACC 3433
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgGly 117
|||||
3434 ATACCGCGGTTGGCCCGTATTAACAGATTACCTCAGGTATCAAGAGG 3483
117 yCysAspValValGluLeu 123
|||||
3484 CTGCGAGCTTGTGCGAGCTC 3502

seq_name: gb_pat:I04278

seq_documentation_block:

LOCUS I04278 1594 bp linear PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent EP 0139417.
ACCESSION I04278
VERSION I04278.1 GI:591810
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 1594)
AUTHORS Berman,P.W. and Lasky,L.A.
TITLE Vaccines based on membrane bound proteins and process for making them
JOURNAL Patent: EP 0139417-A1 2 02-MAY-1985;
FEATURES Location/Qualifiers
source 1.1594
BASE COUNT 310 a 541 c 460 g 283 t
ORIGIN

alignment_scores: Quality: 199.00 Length: 497
 Ratio: 0.948 Gaps: 21
Percent Similarity: 42.254 Percent Identity: 21.932

alignment_block:
US-09-994-064-11 x I04278 ..

Align seg 1/1 to: I04278 from: 1 to: 1594

2 HisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyL 18
:::|||||
54 TACCGACCAACCGACGAAACC.....CCTAAGCGGAGCGC 88
18 lleuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyAlaA 35
:::|:::|
89 GCCATTTTACGAGAGAGAGGGGATATACAAAGTCTGCTTTAAAGCAG 138
35 laValPheThrLeuPheThrTrpThrCysValArgIleMetArgGluHis... 50
:::|
139 GGGTTAGGAGTGTGTCGGTCAT.....AAGCTTCAAGCGGAGACCA 182
51IleCysPheValArgAs 56
:::|
183 ACTACCCGATCATCATGTTATCTTAAGGTCCTTTGTGTGGGCTTC 232
56 naIleMetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleValL 73
:::|
233 CGGTATG..... 239
73 euLeuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
:::|
240GGGGGGCTGCCGCCAGGTTGGGGCCGTGATTTGTGTGTCGTC 284
90 IleLeuGlyArgArg..... 94
:::|
285 ATAGTGGGCTCCATGGGGTCCGGCGCAATATGCCCTTGGCGGATGCCCTC 334
95AlaL 96
:::|
335 TCTCAAGATGGCCGACCCCAATCGCTTTCGGCGCAAGACCTTCGGGTCC 384
96 euAspAlaLeuThrIlePro.....Ala 103
|||||
385 TGGACCACTGACCGACCCCTCCGGGGTCCGGCGGTGTACACATCCAG 434
104 ValGlyProTyrAsnArgTyr.....Le 111
:::|
435 GCGGGCTTACCGAACCCTTCCAGCCCCCAGCCTCCCGATCAGCGTTTA 484
111 uThrArgValSerArgGlyCysAspValValGluLeuAsnProIleSerA 128
|||||
485 CCGCGGTGTGAGCGCGCTGCCGCGCAGCGTCTCTTAACGCAACCGTCGG 534
128 snValAspAspMetIleSerAlaAlaLysGluLysGluLysGlyPro 144
:::|
535 AGGCCCCCAGATTGTCCGGGGGCTCCGAAGACGTCCGGAACAACACC 584
145 PheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspGlyL 161
:::|

```

585  TACAACCTGACCATCGCTTGTTCCGATG.....GGAAG 619
161  uAspLsTYrCYsProIleTYrArgLYsGLuTYrArgLYsGLyAspV 178
    ::      |||||      |||||      |||||      |||||
620  CAACGTGTGCTATCCCAATCACGCGTCATGGAGTACACCGAATGCTCTACA 669
178  aGlnLeuLeuSerGLuCysAlaValGlnSerAlaGlnMetTrpAlaVal 194
    ::      |||:::      ||      :::::      ||
670  ACAAGTCTCTGGGGGCGCTGTCCCATCCGAACGACGCCCGCTGG..... 713
195  AsPTyrValProSer...ThrLeuValSerArgAsnGLyAlaGLyLeuTh 210
    ::|||      ||      ::      |||||:::      ||:::
714  AACTACTATGACAGCTTCAGCGCCGTCAGCGAGATAACCTGGGGTCTCT 763
210  rIlePheSerProThrAlaAlaLeuSerGLyGLnTYrLeuLeuThrLeuL 227
    ::      ::|||:::      ::      ::|||      |||||      ::||
764  GATGCACGCGCCCGCGTTTGAGACCGCGCGCACGTACCTGGCGGCTCGTGA 813
227  ysIleGLyArgPheAlaGlnThrAlaLeuValThrLeuGLu....Val 241
    |||||:::      :::::      ::      |||||      ::
814  AGATTAACGACTGGACGAGATTACACAGTTTATCTCTGGAGCACCGAGCC 863
242  AsnAspArgCYsLeuLYsIleGLySerGlnLeuAsnPhLeuProSerLY 258
    ::      ||      |||      |||||      |||||      |||||
864  AAGGCTCTCTGT...AAGTACACCTCTCCCGCTGCGCATCCCCCGTACG 910
258  sCYsTrpThrThrGLnGLnTYrGLnThrGLy..... 268
    ||      ::      ::      |||||      ||
911  CTGCTCTCTCCCCCAGCGCTACACGAGGGGTGACGGTGACAGCATCG 960
269  .....PheGlnGLyLHisLeuTYrProIleAlaAspThr 280
    |||      |||      |||:::      ::|||
961  GGATGCTGTCCCCGCTTCATCCCCGAGAACCGACCGACCGCTGCGGTATAC 1010
281  AsnThrArgHisAlaAspAspValTYrArgGLyTYrGLuAspIleLeuGL 297
    ::      ::      |||
1011  AGCTTGAAGATCGCCGGG..... 1028
297  nArgTrpAsnAsnLeuLeuArgLYsLYsAsnProSerAlaProAspProA 314
    |||:::      |||||      |||||
1029  ...TGGCAC.....GGGCCAGGGCCCCA..TACA 1053
314  rGProAspSerValProGlnGLnIleProAlaValThrLYsLYsAlaGLu 330
1054  CGAGCACCTGTGCCCCCGGAGCTGCC..... 1082
331  GLyArgThrProAspAlaGLuSerSerGLuLYsLYsAlaProProGLuAs 347
    ::||||||:::      |||||      :::::      ::      |||||
1083  ..GAGACCCCCCAAGGCC.....ACGAGCCAGAACTCGCCCCGGAAGA 1123
347  pSerGLuAspAsp...MetGlnAlaGLuAlaSerGLyGLuAsnProAlaA 363
    |      |||||:::      ::      ::      |||
1124  CCCCAGGATTCGGCCCTCTTGGAAGGACCCCGTGGGACGGTGGCGCCGC 1173
363  lAluProGLuAspAspGLuValProGLuAspThrGLnHisAspAspPro 379
    ::|||      ::      :::::      ||:::      ::      ||||
1174  AAATCCCAACAACGTGGCACTCCCGTCGATCCAGGACGCCCGCGACGCT 1223
380  AsnSerAspProAspTYrTYrAsnAspMetProAlaValIleProValGL 396
1224  TACCATCCCCCGGCACCCCGAACACATGGGCTGATC..... 1262
396  uGLuThrThrLYsSerSerAsnAlaValSerMetProIlePheAlaIap 413
1263  .....GCCGGCGGGGTGGCGGCGAGTCTCTGCGCAGCC 1296
413  heValAlaCysAlaValAlaLeuValGLyLeuLeuValTrp 426
    ::|||      |||      |||:::      |||
1297  TGGTCATTTC.....GGAATTGTGTACTGG 1322
seq_name: gb_pat:I07884
seq_documentation_block:

```

LOCUS	107884	1594 bp	linear	PAT 02-DEC-1994
DEFINITION	Sequence 2 from Patent EP 0139416.			
ACCESSION	107884			
VERSION	107884.1	GI:589401		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1594)			
TITLE	Berman, P.W. and Lasky, L.A.			
JOURNAL	Molecularly cloned diagnostic product and method of use			
FEATURES	Patent: EP 0139416-A2 2 02-MAY-1985;			
source	Location/Qualifiers			
	1..1594			
BASE COUNT	310 a	541 c	460 g	283 t
ORIGIN	/organism="unknown"			

alignment_scores:		
Quality:	199.00	Length: 497
Ratio:	0.948	Gaps: 21
Percent Similarity:	42.254	Percent Identity: 21.932

```
alignment_block:
US-09-994-064-11 x I07884 ...
```

Align seg 1/1 to: I07884 from: 1 to: 1594

[illegible]

```

128 snValaAspMetIleSerAlaAlaLysGluLysGluLysGlyPro 144
    :::: ::::: ::::: ::::: ::::: ::::: :::::
535 AGCCCCCAGATTGTCCGCGGGGCGCTCCGAAGACGTCCGAAACAACC 584
145 pheGluAlaSerValValThrPheTyrValIleLysGlyAspArgLyl 161
    ::::: ::::: ::::: ::::: ::::: :::::
585 TACAACCTGACCATCGCTGTGTTCCGATG.....GGAGG 619
161 uAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGlyAsp 178
    :::: ::::: ::::: ::::: ::::: :::::
620 CAACGTGCTATCCCATCATCGGTACGTAGAGTACACCGAATGCTCTACA 669
178 aGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTrrAlaVal 194
    :::: ::::: ::::: ::::: ::::: :::::
670 ACAAGTCTCTGGGGCGCTGTCCATCCGAACGACGCCCGGCTGG..... 713
195 AspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLeuTh 210
    ::::: ::::: ::::: ::::: ::::: :::::
714 AACTACTATGACAGCTTCAGCGCCGCTCAGCGAGGATAACCTGGGTTCT 763
210 rIlePheSerProThrAlaAlaLeuSerGlyLntYrLeuLeuThrLeuL 227
    ::::: ::::: ::::: ::::: ::::: :::::
764 GATGCACGCCCCCGCTTGTGAGACCGCGGCACGTACCTCGCGCTCTGA 813
227 ysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu.....Val 241
    ::::: ::::: ::::: ::::: ::::: :::::
814 AGATTAACGACTGGACGGAGATTACACAGTTTATCCTGGAGCACCGAGCC 863
242 AsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSerL 258
    ::::: ::::: ::::: ::::: ::::: :::::
864 AAGGCGTCTGT...AAGTACACCTCCCGCTGCGCATCCCCCGCTCAGC 910
258 sCysTrpThrThrGluGlnTyrGlnThrGly..... 268
    ::::: ::::: ::::: ::::: ::::: :::::
911 CTGCTCTCTCCCCCAGCGCTTACACAGGGGGTGACGGTGACACGATCG 960
269 .....PheGlnGlyGlnHisLeuTyrProIleAlaAspThr 280
    ::::: ::::: ::::: ::::: ::::: :::::
961 GGATGCTGCCCCGCTTCATCCCGGAGAACACGACCGACCGCTGCGCTATAC 1010
281 AsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLeuG 297
    ::::: ::::: ::::: ::::: ::::: :::::
1011 AGCTTGAAGATCGCCGG..... 1028
297 nArgTrrAsnAsnLeuLeuArgLysLysAsnProSerAlaProAspPro 314
    ::::: ::::: ::::: ::::: ::::: :::::
1029 ....TGGCAC.....GGGCCAGGGCCCA...TACA 1053
314 rGProAspSerValProGlnGluIleProAlaValThrLysLysAlaGlu 330
    ::::: ::::: ::::: ::::: ::::: :::::
1054 CGAGCACCTGTGCTGCCCCCGAGCTGCC..... 1082
331 GlyArgThrProAspAlaGlnSerSerGluLysLysAlaProProGluAs 347
    ::::: ::::: ::::: ::::: ::::: :::::
1083 ...GAGACCCCCCAACGCC....ACGCAGCCAGAACTGCCCCCGGAAGA 1123
347 pSerGluAspAsp...MetGlnAlaGluAlaSerGlyLysAsnProAlaA 363
    ::::: ::::: ::::: ::::: ::::: :::::
1124 CCCCAGGATTCGGCCCTTGTGAGGAGACCCCGTGGGACGGTGGCGCCG 1173
363 IalauProGluAspAspArgLysValProGluAspThrGluHisAspAsp 379
    ::::: ::::: ::::: ::::: ::::: :::::
1174 AAATCCCAACAACCTGCGACATCCCGTCGATCCAGGACGCCGCGACGCT 1223
380 AsnSerAspProAspTyrTrrAsnAspMetProAlaValIleProValG 396
    ::::: ::::: ::::: ::::: ::::: :::::
1224 TACCATCCCCCGGCCACCCCGAACACATGGGCGCTGATC..... 1262
396 uGluThrThrLysSerSerAsnAlaValSerMetProIlePheAlaAla 413
    ::::: ::::: ::::: ::::: ::::: :::::
1263 .....GCCGGCGCGGTGGCGGCGGACGTCTCTGCGACGCC 1296
413 heValAlaCysAlaValAlaLeuValGlyLeuLeuValTrr 426

```

```

seq_name: gb_vi:HSIGDB
seq_documentation_block:
LOCUS      HSIGDB              1594 bp      DNA      linear   VRL 02-AUG-1993
DEFINITION Herpes simplex virus type 1 glycoprotein-D gene, complete cds.
ACCESSION  K02372
VERSION    K02372.1  GI:330102
KEYWORDS   glycoprotein; glycoprotein D.
SOURCE     HSV type 1 (strain Hzt) DNA.
ORGANISM   human herpesvirus 1
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Simplexvirus.
REFERENCE  1 (bases 1 to 1594)
            Lasky, L.A. and Dowdenko, D.J.
            DNA sequence analysis of the type-common glycoprotein-D genes of
            herpes simplex virus types 1 and 2
            DNA 3, 23-29 (1984)
JOURNAL    84131549
MEDLINE
FEATURES
            source
            location/Qualifiers
                1. 1594
                    /organism="human herpesvirus 1"
                    /db_xref="taxon:10298"
                    237. 1418
                    /note="glycoprotein-D"
                    /codon_start=1
                    /protein_id="AAA45786.1"
                    /db_xref="GI:330103"
                    /translation="MGGAARLGAVILFVIVIGLHGVRGYALADSLKADPNRERG
                    KDLPLYDLDTDPBGRVRRVYHIQAGLPNRPQPSLPITVYRVRACRSVLLNAPSEAP
                    QIVRGASEDVARKQPNLTIAWFRMGNCAPITVMEYTECSYNSLSGACPRTQPRWN
                    YYDSFSAVEDNLGFLMHARAFETAGYLRKINDWTEITQFILEHRAKGSCKYTLR
                    LRIPPSACLSPOAYOOQGVTVDSIGMLPRIPENORTVAVSLKIAGHGRAPYTSLL
                    LPPELPETPNATQPELAPDEPDSDLEDPVGTAVARQIPRNWHIPSTQDAATPYHPPA
                    TPNNMGLIAGAVGSSLALALVICGIYVWMRRTRKAPKRIRLPRIREDDPSSHQPLF
                    Y"
BASE COUNT  310 a      541 c      460 g      283 t
ORIGIN
alignment_scores:
            quality: 199.00      Length: 497
            Ratio: 0.948      Gaps: 21
            Percent Similarity: 42.254      Percent Identity: 21.932
alignment_block:
US-09-994-064-11 x HSIGDB      ..
Align seg 1/1 to: HSIGDB from: 1 to: 1594
2  HisArgProHisLeuArgArgHisSerArgTyrTyraAlaLysGlyGluVa 18
   ::::::::::::::  |||  ::::::::::::::
54  TACCGACACACACCGACGAACC.....CCPAAGGGGGAGGG 88
18  IleuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyAlaA 35
   ::::  ::::  ::::  ::::  |||  ::::  ::::  ::::
89  GCCATTTTACGAGAGGAGGGGTATACAAAGTCTGCTTTAAAAAAGCAG 138
35  laValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis... 50
   ::|||  |||||  ::::  ::::  ::::  ::::  ::::
139  GGGTTAGGAGTGTGTCGGTCAT.....AAGCTTCAGCGCGAAGACCA 182
51  .....IleCysPheValArgAs 56
183  ACTACCCCGATCATCATGATATCCTTAAGGTCCTTTTGTGTGTCGCCGTC 232
56  naIaMetAspArgHisLeuPheLeuArgAsnaIaPheTrpThrIleValL 73
   ::::|||
233  CGGTATG..... 239
73  euleuSerSerPheAlaSerGlnSerThrAlaAlaValIThrTyrAspTyr 89

```

```

240 .....GGGGGGCTGCCGACGTTGGGGCCCGTGATTTGTTGTCGTC 284
      ::::: |||::: |||:::
90 IleuGlyArgArg..... 94
      ||::: ||| :::
285 ATAGTGGGCTCCATGGGGTCCGGCGCAATATGCCTTGGCGGATGCCCTC 334
95 .....AlaL 96
      ::::
335 TCTCAAGATGGCCGACCCCAATCGCTTCGGGCAAGACCTTCGGGTCC 384
96 euAspAlaLeuThrIlePro.....Ala 103
      ||||| ||||| |||
385 TGGACCACTGACCGAACCTCCGGGGTCCGGCGGTGATCCACATCCAG 434
104 ValGlyProTyrAsnArgTyr.....Le 111
      :::: ||| :::
435 GCGGCGCTACCGAACCCGTTCCAGCCCCCAGCCTCCGATCAGGTTTA 484
111 uThrArgValSerArgGlyCysAspValValGluLeuAsnProIleSerA 128
      |||||::: ||| ||||| |||
485 CCGCGGTGTGGAGCGGCTGCCGACGCTGCTCTAAACGACACCGTCGG 534
128 snValAspMetIleSerAlaAlaLysGluLysGluLysGlyGlyPro 144
      :::: ::::: ::::: |||::: |||
535 AGGCCCCCAGATGTCCGGGGGCTCCGAGAGACGTCCGAAACAAACCC 584
145 pheGluAlaSerValValTrpPheTyrValIleLysGlyAspArgLys 161
      ::::: ::::: ||||| ::: |||
585 TACAACCTGACCAATCGCTGTGTTCCGATG.....GGAGG 619
161 uAspLysTyrCysProIleTyrArgLysGluTyrArgLysGlyAsp 178
      ::: ||||| ||||| ||||| :::
620 CAACTGTGCTATCCCATCAGGTCATGAGTACACCGAATGCTCCTACA 669
178 alGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTrpAlaVal 194
      ::: ||::: ||| ::::: |||
670 ACAAGTCTCTGGGGCTGTCCCATCCGAACGACGCCCCGCTGG..... 713
195 AspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLeuTh 210
      :::: ||| ::: |||||::: ||:::
714 AACTACTATGACAGCTTCAGCGCCGTCAGCGAGGATTAACCTGGGGTCT 763
210 rIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrLeu 227
      ::: :::: ||::: ||||| ::::
764 GATGCACGCCCCGCTTGAAGACCGCCGACGTAACCTGCGGCTCGTGA 813
227 ysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu.....Val 241
      |||||::: ::::: ::: |||||
814 AGATAAACGACTGACGCGAGATTACACAGTTTATCTTGGAGACACCGAGCC 863
242 AsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSerLys 258
      ::: ||| ||| |||::: |||||
864 AAGGGCTCCTGT...AAGTACACCCCTCCCGCTGCCCATCCCCCGTCAGC 910
258 scysTrpThrThrGluGlnTyrGlnThrGly..... 268
      ||| ::: ||||| |||
911 CTGCCCTCTCCCCCAGGCTTACACAGAGGGGTGACGCTGACAGCATCG 960
269 .....PheGlnGlyGluHisLeuTyrProIleAlaAspThr 280
      ||| |||::: :::: |||
961 GGATGCTGCCCGCTTATCCCGAGAACACGACGCGCATCGCGGTATAC 1010
281 AsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLeuG 297
      ::: ::: |||
1011 AGCTTGAAGATCGCCGG..... 1028
297 nArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAspProA 314
      ||::: ||::: |||||
1029 .....TGGCAC.....GGGCCAGGGCCCA...TACA 1053
314 rgProAspSerValProGlnGluIleProAlaValThrLysLysAlaGlu 330
      :::: ||| ||::: |||
```

```

1054 CGAGCACCCCTGCTGCCCGGAGCTGCC..... 1082
331 GlyArgThrProAspAlaGluSerSerGluLysLysAlaProProGluAs 347
      :::: |||||::: ||| ::::
1083 ...GAGACCCCCCAACGCC.....ACGACGCCAGAACTCGCCCGGAAGA 1123
347 pSerGluAspAsp...MetGlnAlaGluAlaSerGlyGluAsnProAlaA 363
      | |||||::: ::: |||
1124 CCCCAGAGATTCGGCCCTTGGAGAGACCCCGTGGGACGGTGGCGCCGC 1173
363 lAlaLeuProGluAspAspGluValProGluAspThrGluHisAspAspPro 379
      :::: ||| ::: |||||:::
1174 AAATCCCAACCAACTGGACATCCCGTCGATCCAGAGACGCCGACGCT 1223
380 AsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProValG 396
      ||| |||::: ||| :::
1224 TACCATCCCCCGCCACCCCGAACACATGGGCGCTGATC..... 1262
396 uGluThrThrLysSerSerAsnAlaValSerMetProIlePheAlaAla 413
1263 .....GCCGGCGGCTGGCGGCGCAGTCTCTGCGACAGCCC 1296
413 heValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
      ::||| ||| |||::: |||
1297 TGGTCATTTGC.....GGAATTGTGTACTGG 1322

seq_name: gb_pat:E00395
seq_documentation_block:
LOCUS E00395 1594 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA coding herpes simplex virus 1 gd (HSV-1 gd).
ACCESSION E00395
VERSION E00395.1 GI:2168680
KEYWORDS JP 198515128-A/2.
SOURCE human herpesvirus 1.
ORGANISM human herpesvirus 1.
          viruses; dsDNA viruses, no RNA stage; Herpesviridae;
          Alphaherpesvirinae; Simplexvirus.
REFERENCE 1 (bases 1 to 1594)
AUTHORS Fuiritsupu,U.B. and Roorensu,A.R.
TITLE VACCINES CONTAINING MEMBRANE BINDABLE PROTEIN
JOURNAL Patent: JP 198515128-A 2 15-AUG-1985;
          GENENTECH INC
COMMENT OS human herpesvirus 1
        PN JP 198515128-A/2
        PD 15-AUG-1985
        PF 30-AUG-1984 JP 1984183623
        PR 30-AUG-1983 US 83 527917, 31-OCT-1983 US 83 547551, PR
        PI FUIRITSUPU UEIN BAAMAN, ROORENSU ARAN RASUKII PC
        A61K39/245,C12N15/00//C12P21/00;
        CC strandedness: Double;
        CC topology: Linear;
        CC hypothetical: No;
        CC anti-sense: No;
        CC Feature is identified by experimental;
        FH Key location/Qualifiers
        FH mat_peptide 237..1416
        FT /product='HSV-1 gd'
        FT CDS 237..1419
        FT /product='HSV-1 gd'
        FT TATA_signal 167..171
        FT /number=1
        FT TATA_signal 186..191
        FT /number=2
        FT TATA_signal 1537..1541
        FT /number=3.
FEATURES
          source location/Qualifiers
          1..1594
          /organism="human herpesvirus 1"
          /db_xref="taxon:10298"
BASE COUNT 308 a 542 c 461 g 283 t
```


ORIGIN

alignment_scores:

Quality: 198.00 Length: 342
Ratio: 1.215 Gaps: 15
Percent Similarity: 47.661 Percent Identity: 24.269

alignment_block:

US-09-994-064-11 x E00395 ..

Align seg 1/1 to: E00395 from: 1 to: 1594

101 ILeProAlaValGlyProTyrAsnArgTyrLeu.ThrArgValSerArg 117
:::||||| ||| :::: ||| |||||:::|
452 GTTCAGACCCCAAGCTCCCGATCACGGTTACCGCGGTGGAGCGGG 501
117 LcYAspValValGluLeuAsnProIleSerAsnValAspMetIle 133
::||| ||| ||||| ||:::|
502 CCTGCCGACGCGTCTTAACGACCGCTCGAGGCCCGCCAGATTGTC 551
134 SerAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVa 150
:::||||| ||| :::: ||| |||||:::|
552 CGCGGGCGCTCCGAGACGTCGGAACAACCTTACACCTGACCATCGC 601
150 lTrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProI 167
:||||| :::: ||| :::: ||| |||||
602 TTGGTTTCGGATG.....GGAGCAACTGTGTATATCCCA 636
167 lETyrArgLysGluTyrArgGluCysGlyAspValGluLeuLeuSerGlu 183
|| ||||| |||||::: ||::|
637 TCACGTCATGAGTACACCGAATGCTCTACACAAGTCTTGGGGCC 686
184 CysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSer.. 199
||| ::::| ||| ::::| |||
687 TGTCCATCCGACGACGACCGCGCTG.....AACTACTATGACAGCTT 730
200 .ThrLeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThra 216
::: |||||:::| ||::: ||::|
731 CAGCGCGCTCAGCGAGGATTAACCTGGGTTCTGTATGACGCGCGCGT 780
216 lAlaLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAla 232
:::| ||| ||||| ::::|
781 TTGAGACCGCGCGACGTACCTGCGGCTGTGAGATTAACGACTGACG 830
233 GlnThrAlaLeuValThrLeuGlu.....ValAsnAspArgCysLeuTy 247
::: ||: ||||| ::::| ||| ||
831 GAGATTACACAGTTTATCTGTGAGCACCAGCCAGGGCTCTGT..AA 877
247 sIleGlySerGlnLeuAsnPheLeuProSerLysCysTrpThrThrGluG 264
| ||:::| ||||| ||| ::::|
878 GTACACCTCCCGCTGCGCATCCCCCGTCAGCCTGCTCTCCCCCAGG 927
264 lntYrGlnThrGly.....Phe 269
||| ||| |||
928 CCTACCAGCAGGGGGTGACGGTGACAGCATCGGATGCTGCCCGCTTC 977
270 GlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgHisAlaAs 286
||:::| ||:::| ||:::|
978 ATCCCCGAGAACGACGACCGCTGCGGTATACAGCTTGAAGATCGCCG 1027
286 pAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsnAsnLeuL 303
||:::|
1028 G.....TGGCAC..... 1034
303 euArgLysLysAsnProSerAlaProAspProArgProAspSerValPro 319
:::| ||| ||| |||
1035GGGCCAGGGCCCA..TACACGAGACACCTGCTGCC 1070
320 GlnGluIleProAlaValThrLysLysAlaGluGlyArgThrProAspAl 336
||:::| ||| ||| |||
1071 CCGAGAGCTGCC.....GAGACCCCAACGC 1096

336 aGluSerSerGluLysLysAlaProProGluAspSerGluAspAsp...M 352
| ::::| ||||| |||||:::|
1097 C.....ACGACACCACTCGCCCGGAGAGACCCGAGATTTCGCCCC 1140
352 etGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAspAsp 368
::: ||: ||| ::::| |||
1141 TCTTGAGGAGACCCCGTGGGACGGTGCGCCCGCAATCCCAACCACTGG 1190
369 GluValProGluAspThrGluHisAspAspProAsnSerAspProAspTy 385
:::| ||:::| ||| |||
1191 CACATCCCGTGCATCGAGGACGCGCGACGCGCTTACCATCCCCCGCCAC 1240
1241 CCCGAACAACATGGCGCTGATC.....G 1263
402 erAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAlaVal 418
:::| |||||:::| ::::| ||||| |||
1264 CCGCGCGGTGGCGGCGAGTCTCTGCGACGCCCTGTCATTTC..... 1307
419 AlaLeuValGlyLeuValTrp 426
||:::| |||
1308GGAATTGTGTACTGG 1322

seq_name: gb_pat:E00401

seq_documentation_block:

LOCUS E00401 1594 bp DNA linear PAT 29-SEP-1997
DEFINITION Herpes simplex virus-1 (HSV-1) glycoprotein D (gD) gene.
ACCESSION E00401
VERSION E00401.1 GI:2168684
KEYWORDS JP 1985155974-A/2.
SOURCE human herpesvirus 1.
ORGANISM human herpesvirus 1.
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.

REFERENCE
AUTHORS Fuiritsupu,U.B. and Roorensu,A.R.
TITLE DIAGNOSTIC PRODUCT CHANGED INFO MOLECULAR CLONE
JOURNAL Patent: JP 1985155974-A 2 16-AUG-1985;
GENENTECH INC

COMMENT
OS herpes simplex virus type 1
PN JP 1985155974-A/2
PD 16-AUG-1985
PF 30-AUG-1984 JP 1984183622
PR 30-AUG-1983 US 83 527916, 31-OCT-1983 US 83 547552, PR
09-MAR-1984 US 84 587763
PI FUIRITSUPU UEIN BAAMAN, ROORENSU ARAN RASUKII PC
G01N33/54,A61K39/00,C07K15/14,C12N15/00,G01N33/577//A61K49/00, PC
C12P21/00;
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Hzt strain;
CC Feature is identified by experimental;
FH Key location/Qualifiers
FT TATA_signal 167..171
FT FT /number=1
FT TATA_signal 186..191
FT FT /number=2
FT TATA_signal 1537..1541
FT FT /number=3
FT mat_peptide 237..1416
FT FT /product='HSV-1 gD'
FT FT 237..1419
FT CDS /product='HSV-1 gD'.
FT FT location/Qualifiers
FT 1..1594
FT /organism='human herpesvirus 1'
FT /db_xref='taxon:10298'

BASE COUNT 308 a 542 c 461 g 283 t


```
187 ACTACCCGATCATGACTTATCCTTAAGTCTCTTTTGTGTGTCGCTTC 236
56 nAlaMetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleValL 73
   ::::|
237 CGGTATG..... 243
73 euLeuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
   ::::|
244 ....GGGGGACTGCCGCCAGGTTGGGGGCCGTGATTGTTGTGTCGTC 288
90 IleLeuGlyArgArg..... 94
   ::::|
289 ATAGTGGCCTTCATGGGGTCCGGCGCAATATGCTTGCGGATGCCCTC 338
95 .....AlaL 96
339 TCTCAAGATGCCGACCCCAATCGCTTTCGCGCAAGACCTTCGCGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
   ::::|
389 TGGACCAAGCTGACCGACCTCCGGGGTCCGGCGGTGTACACATCCAG 438
104 ValGly.....ProTyrAsnArgTyr 110
   ::::|
439 GCGGGCCTACCGGACCCGTTCCAGCCCCCAGCCCTCCCGATCAGGTTTA 488
110 rLeuThrArgValSerArgGlyCysAspValValGluLeuAsnProIleS 127
   ::::|
489 CTACGCCGTGTGGAGCGCGCCTGCGCGACGCTCTCTAAACGACACCGT 538
127 eRAsnValAspAspMetIleSerAlaAlaLysGluLysGluLysGly 143
   ::::|
539 CGGAGGCCCCCGCATTTGTCGGCGGCTCCGAAGACGTCCGGAACAA 588
144 PropheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspG 160
   ::::|
589 CCCTACAACTGACCATCGCTTGGTTCCGATG.....GG 623
160 yGluAspLysTyrCysProIleTyrArgLysGluTyrArgLysGlyA 177
   ::::|
624 AGCAACTGTGCTATCCCCATCAGGTCATGAGTACACCGAATGCTCCT 673
177 spValGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTrpAla 193
   ::::|
674 ACAACAAGTCTTGGGGGCTGTCCCATCCGAACGACGCCCGCTGG... 720
194 ValAspTyrValProSer..ThrLeuValSerArgAsnGlyAlaGlyLe 209
   ::::|
721 ..AACTACTATGACAGCTTACGCGCGCTCAGCGAGGATTAACCTGGGTT 767
209 uThrIlePheSerProThrAlaAlaLeuSerGluGlnTyrLeuLeuThrL 226
   ::::|
768 GCTGATGCACGCCCGCGCTTGGAGACCGCGCACGTACCTGCGGCTCG 817
226 euLysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
   ::::|
818 TGAAGATAAACGACTGGACGGAGATTACACAGTTTATCTGGAGCACCGA 867
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSe 257
   ::::|
868 GCCAAGGGCTCTGT..AAGTACGCCCTCCCGCTGCGCATCCCCCGCTC 914
257 rLysCysTrpTrpThrThrGluGlnTyrGlnThrGly..... 268
   ::::|
915 AGCCTGCTCTCCCCCGCTTACACAGGGGGTGCAGGTTGACAGCA 964
269 .....PheGlnGlyLuhHisLeuTyrProIleAlaAsp 279
   ::::|
965 TCGGATGCTGCCCGCTTCATCCCCGAGAACACGACGCTCGCCGTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
   ::::|
1015 TACAGCTTGAAGATCGCCGGG..... 1035
```

```
296 uGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
   ::::|
1036 .....TGGCAC.....GGGCCCAAGGCCCA...T 1057
313 roArgProAspSerValProGlnLuhIleProAlaValThrLysLysAla 329
   ::::|
1058 ACACGAGCACCTGCTGCCCGCGAGCTG..... 1086
330 GluGlyArgThrProAspAlaGluSerSerGluLysLysAlaProProG 346
   ::::|
1087 ...TCCGAGACCCCAACGCC....ACGACGCCAGAACTCGCCCCGA 1127
346 uAspSerGluAspAsp...MetGlnAlaGluAlaSerGlyLuhAsnProA 362
   ::::|
1128 AGACCCCGAGGATTCGGCCCTCTTGAGAGACCCCGTGGGACGTTGCCG 1177
362 lAlaLeuProGluAspAspGluValProGluAspThrGluHisAspAsp 378
   ::::|
1178 CGCAATCCCAACCAACTGGACACATCCCGTCGATCCAGGACGCCGCGAC 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProva 395
   ::::|
1228 CCTTACCATCCCCCGCGCACCCCGAACACATGCGCTGATC..... 1269
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAla 412
   ::::|
1270 .....GCCGGCGGTGGCGGCGCAGTCTCCTGCGCAG 1300
412 lApeValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
   ::::|
1301 CCCTGGTCATTTGC.....GGAATTGTGTACTGG 1329
seq_name: gb_pat:AR043518
seq_documentation_block:
LOCUS AR043518 1608 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5814486.
ACCESSION AR043518
VERSION AR043518.1 GI:5964526
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
BASE COUNT 313 a 542 c 465 g 288 t
ORIGIN
alignment_scores:
Quality: 193.50 Length: 498
Ratio: 0.908 Gaps: 21
Percent Similarity: 42.771 Percent Identity: 21.486
alignment_block:
US-09-994-064-11 x AR043518 ..
Align seg 1/1 to: AR043518 from: 1 to: 1608
2 HisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyGluVa 18
   ::::|
58 TACCGACACACCGACGAAC.....CTAAGGGGAGGAGG 92
18 lLeuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyAla 35
   ::::|
93 GCCATTTTACGAGGAGGAGGGGTATTAACAAGTCTGTCTTAAAAAGCAG 142
```

```
35 laValPheThrIleuPheTrpThrCysValAlaArgIleMetArgGluHis... 50
   ::||| ||||| :::::|||||:::
143 GGGTAGGAGTTGTCGGTCAAT...AAGCTTCAGCGCGAAGCAGCA 186
51 .....IleCysPheValArgAs 56
187 ACTACCCCGATCATCAGTTATCCTTAAGTCTCTTTGTGTGGTGGCTTC 236
56 nAlaMetAspArgHisIleuPheLeuArgAsnAlaPheTrpThrIleValL 73
   ::::|||
237 CGGTATG..... 243
73 euLeuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
   ::::: |||::: |||::: :::
244 ....GGGGGACTGCGCCGACAGGTGGGGCGCGTGATTTGTGTGTCGTC 288
90 IleuGlyArgArg..... 94
289 ATAGTGGGCTCCATGGGGTCCGCGCAATATGCTTGCGGAGTGCCTC 338
95 .....AlaL 96
339 TCTCAAGATGGCCGACCCCAATCGCTTTCCGGGCAAGACCTCCGGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
389 TGGACCACTGACCGACCCCTCGGGGGTCCGGCGGTATCACATCCAG 438
104 ValGly.....ProTyrAsnArgTy 110
439 GCGGGCTTACCGGACCCGTTCCAGCCCCCAAGCCTCCCGATCACGGTTTA 488
110 rLeuThrArgValSerArgGlyCysAspValValGluLeuAsnProIles 127
489 CTACGCGTGTGTGGAGCGCGCTGCCGCGACGCTGCTCTTAACGCACCGT 538
127 eAsnValAspAspMetIleSerAlaAlaLysGluLysGluLysGly 143
539 CGGAGGCCCCCAAGATTGTCCGCGGGCCCTCCGAAGACGTCGGAACAA 588
144 ProPheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspG 160
589 CCCTACAACTGACCATCGCTTGCTTTCGGATG.....GG 623
160 yGluAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGlyA 177
624 AGGCAACTGTGCTATCCCATCAAGGTCATGAGTACACCGAATGCTCCT 673
177 spValGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTrpAla 193
674 ACAACAAGTCTCTGGGGCTGTCCCATCCGAACGACGCCCGCTGG... 720
194 ValAspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLe 209
721 ..AACTACTATGACAGCTTCAGCGCGCTCAGCGAGGATTAACCTGGGCTT 767
209 uThrIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrL 226
768 CCTGATGCACGCCCCCGCTTGAAGACCGCGCGACGTAACCTGGGCTCG 817
226 euLysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
818 TGAAGATAAACGACTGGACGGAGATTACACAGTTTATCTTGAAGACACCGA 867
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProse 257
868 GCCAAGGGCTCCTGT...AAGTACGCCCTCCCGCTGCGCATCCCCCGTGC 914
257 rLysCysTrpThrThrGluGlnTyrGlnThrGly..... 268
915 AGCCTGCTCTCCCGCCAGGCTTACACAGCAGGGGTGACGGTGACAGCA 964
269 .....PheGlnGlyGluHisLeuTyrProIleAlaAsp 279
```

```
965 TCGGGATGCTGCCCGGCTTATCCCGGAGAACGACGCGACCGTCCGCTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
1015 TACAGCTTGAAGATCGCCGG..... 1035
296 uGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
1036 .....TGGCAC.....GGCCCAAGCCCCCA...T 1057
313 roArgProAspSerValProGlnGluIleProAlaValThrLysLysAla 329
1058 ACACGAGCACCCCTGCTGCCCCCGAGCTG..... 1086
330 GluGlyArgThrProAspAlaGlnSerSerGluLysLysAlaProProG 346
1087 ...TCCGAGACCCCAACGCC....ACGACGCCAGAACTCGCCCCGGA 1127
346 uAspSerGluAspAsp...MetGlnAlaGluAlaSerGlyGluAsnProA 362
1128 AGACCCCGAGGATTCGGCCCTCTTGAGAGACCCCGTGGGGACGGTGGCC 1177
362 IaAlaLeuProGluAspAspGluValProGluAspThrGluHisAspAsp 378
1178 CGCAATCCACCAACTGGACATCCCGTGCATCCAGGACGCGCGGACG 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProva 395
1228 CTTTACCATCCCCCGGCCACCCCGAACAACATGGGCTGATC..... 1269
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAlaA 412
1270 .....GCCGGCGGGTGGCGGACGTCTCTCGGCAG 1300
412 lApheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
1301 CCCTGTCATTTGC.....GGAATGTGTACTGG 1329

seq_name: gb_pat:E00204
seq_documentation_block: 1608 bp DNA linear PAT 29-SEP-1997
LOCUS E00204
DEFINITION DNA sequence encoding gd-1 of herpes simplex virus typel.
ACCESSION E00204
VERSION E00204.1 GI:2168500
KEYWORDS JP 1984118097-A/1.
SOURCE human herpesvirus 1.
ORGANISM human herpesvirus 1.
VIRUSES: dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
REFERENCE 1 (bases 1 to 1608)
AUTHORS Rojiyaa,J.W., Jiyon,H.W. and Rin,U.E.
TITLE PRODUCTION OF HERPES SIMPLEX VIRUS PROTEIN
JOURNAL Patent: JP 1984118097-A 1 07-JUL-1984;
MORUKIYURAA JIENETEITSUKUSU INC
COMMENT OS Herpes simplex virus type 1
PN JP 1984118097-A/1
PD 07-JUL-1984
PF 20-JUL-1983 JP 1983131151
PR 20-JUL-1982 US 82 400028, 25-OCT-1982 US 82 436368 PI
ROJIYAA JIYON WATOSON, JIYON HEINZU WAISU,
PI RIN URIAMU ENKISUTO
PC C12P21/00,A61K39/245,C07G7/00,C07H21/04,C12N1/00,C12N15/00,PC
(C12P21/00,
PC C12R1:19),(C12N1/00,C12R1:19),(C12N1/00,C12R1:19); CC
strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone-prwF6;
FH Key Location/Qualifiers
```


FT 5'UTR 1. .240
FT CDS 241. .1425
FT /product='gd-1 of herpes simplex virus type
FT 1

FEATURES
source location/Qualifiers
1. .1608

BASE COUNT 313 a 543 c 464 g 288 t
/db_xref="taxon:10298"

ORIGIN

alignment_scores:

Quality: 193.50 Length: 498
Ratio: 0.908 Gaps: 21
Percent Similarity: 42.771 Percent Identity: 21.486

alignment_block:

us-09-994-064-11 x E00204 ..

Align seg 1/1 to: E00204 from: 1 to: 1608

2 HisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyGluVa 18
:::|||||::|
58 TACCGACCACACCGACGAACC.....CCTAAGGGGAGGG 92
18 IleAsnLysHisMetAspCysGlyLysArgCysCysSerGlyAlaA 35
:::|::|::|
93 GCCATTTTACGAGGAGGAGGGGTATACAAAGTCTCTTTAAAAACAG 142
35 IValAlpheThrLeuPheTrpThrCysValArgIleMetArgLuhis... 50
:::|::|::|
143 GGGTTAGGAGTTGTCGTCA.....AAGCTTCAGCGCGAAGCAGCA 186
51IleCysPheValArgAs 56
187 ACTACCCCGATCATCAGTTATCCTTAAGTCTCTTTGTGTGGTGGCTC 236
56 nAlaMetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleVal 73
:::|::|
237 CGGTATG..... 243
73 euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
:::|::|::|
244GGGGGACTGCCGCGCAGGTGGGGGCGCGTGAATTTGTTGTCGTC 288
90 IleuGlyArgArg..... 94
289 ATAGTGGCCTCCATGGGGTCCGCGCAATATGCTTGGCGGATGCCCTC 338
95AlaL 96
339 TCTCAAGATGGCCGACCCCAATCGCTTTCGCGCAAGACCTCCGGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
|||||::|
389 TGGACCACTGACCGACCCCTCCGGGGGTCCGGCGCGTGTACCATCCAG 438
104 ValGly.....ProTyrAsnArgTyr 110
:::|::|
439 GCGGGCTTACCGGACCCGTTCCAGCCCCCAGCCTCCCGATCACGGTTTA 488
110 rLeuThrArgValSerArgGlyCysAspValValGluLeuAsnProIles 127
|:::|::|::|
489 CTACGCCGTGTGGAGCGCGCTGCCGACGCGTGTCTTAACGACCGCT 538
127 eAsnValAspAspMetIleSerAlaAlaLysGluLysGlyGlyGly 143
|:::|::|::|
539 CGAGGCCCCCAGATTGTCCGCGGGCTCCGAGAGACGTCCGGAAACAA 588
144 ProPheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspG 160
|:::|::|::|
589 CCCTACAACTGACCATCGTTGTTTCGGATG.....GG 623

160 yGluAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGlyA 177
|:::|::|
624 AGGCAACTGTGCTATCCCATCATCAGCGGTACAGTAGTACACCGAATGCTCT 673
177 spValGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetPala 193
:::|::|::|
674 ACACAAGTCTTGGGGGCGCTGCCATCCGAACGACGACCCCGCTGG... 720
194 ValAspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLe 209
:::|::|
721 ...AACTACTATGACAGCTTACAGCGCGGTACAGGATTAACCTGGGGTT 767
209 uThrIlePheSerProThrAlaAlaLeuSerGlyLysTyrLeuLeuThrL 226
:::|::|
768 CCTGATGCACGCGCGCGCTTGAGACCGCGCGCACGTACCTGCGGCTCG 817
226 eulLysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
:::|::|
818 TGAAGATAACGACTGACGAGGAGATTACAGATTATCTTGAGACACCGA 867
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSe 257
:::|::|
868 GCCAAGGGCTCTGT...AAGTACGCCCTGCCGCTGCGCATCCCGCTGC 914
257 rLysCysTrpThrThrGluGlnTyrGlnThrGly..... 268
|::|
915 AGCCTGCCCTCCTCCCGCCAGGCTTACACAGAGGGGTGACGGTGGACAGCA 964
269PheGlnGlyLuhisLeuTyrProIleAlaAsp 279
|::|
965 TCGGATGCTGCGCGCGCTTCATCCCGAGAACACGCGACCGTCCGCCGTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
:::|::|
1015 TACAGCTTGAAGATCGCCGGG..... 1035
296 uGlnArgTrpAsnAsnLeuLeuArgLysAsnProSerAlaProAsp 313
|::|
1036TGGCAC.....GGGCCAAGGCCCA...T 1057
313 roArgProAspSerValProGlnGluIleProAlaValThrLysLysAla 329
:::|::|
1058 ACACGAGCACCTGCTGCCCGCCGAGCTG..... 1086
330 GluGlyArgThrProAspAlaGluSerSerGluLysLysAlaProProG 346
:::|::|
1087 ...TCCGAGACCCCAACGCG...ACGACGCCAGAACTCGCCCGCGA 1127
346 uAspSerGluAspAsp...MetGlnAlaGluAlaSerGlyGluAsnProA 362
|::|
1128 AGACCCCGAGGATTCGGCCCTCTTGGAGAGACCCCGTGGGACGGTGGCGC 1177
362 IaAlaLeuProGluAspAspGluValProGluAspThrGluHisAspAsp 378
:::|::|
1178 CGCAAAATCCACCAAACTGGCACATCCGTCGATCCAGAGCGCGGACG 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProVa 395
|::|
1228 CCTTACCAATCCCCCGGCGACCCCGAACAACATGGCGCTGATC..... 1269
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAlaA 412
:::|::|
1270GCCGGCGGGTGGCGGACGTCTCTGGCAG 1300
412 lApheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
|::|
1301 CCCTGCTCATTTGC.....GGAATTGTGTACTGG 1329

seq_name: gb_pat:101122

seq_documentation_block:

LOCUS I01122 1608 bp ss-DNA linear PAT 21-MAY-1993

```

DEFINITION   Sequence 1 from Patent US 4762779.
ACCESSION    I01122
VERSION      I01122.1  GI:269313
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 1608)
AUTHORS      Snitman,D.L.
TITLE        Compositions and methods for functionalizing nucleic acids
JOURNAL      Patent: US 4762779-A 1 09-AUG-1988;
              Amgen Inc.;
              Thousand Oaks, CA

FEATURES
  source          Location/Qualifiers
                   1..1608
BASE COUNT      313 a      542 c      465 g      288 t
ORIGIN
alignment_scores:
  Quality:      193.50      Length:      498
  Ratio:        0.908      Gaps:      21
  Percent Similarity: 42.771      Percent Identity: 21.486

alignment_block:
  US-09-994-064-11 x I01122      ..

Align seg 1/1  to: I01122  from: 1  to: 1608

      2  HISARGPROHISLEUARGHISERARGTYRTRALALYSGLYGUVA 18
          :::::::::::::: ::
      58  TACCGACACACACCGACGAAACC.....CCTAAGGGGAGGG 92
          :::::::::::::: ::
      18  ILEUASNLYSHISMETASPCYSGLYLYSARGCYSCYSSERGLYALA 35
          :::: :::: :::: :::: ::::
      93  GCCATTTTACGAGGAGGAGGGGTATACAAAGTGTCTTTAAAGAAGCAG 142
          :::: :::: :::: :::: ::::
      35  IAVAlPheThrLeuPheTrpThrCysValArgIleMetArgGluHis... 50
          ::::: ::::: ::::: ::::: :::::
      143  GGGTTAGGAGTGTGTCGGTCAT....AAGCTTCAGCGCGAAGCAGCA 186
          ::::: ::::: ::::: ::::: :::::
      51  .....IleCysPheValArgAs 56
          ::::: ::::: ::::: ::::: :::::
      187  ACACACCCGATCATCAGTTATCCTTAAGGTCTCTTTGTGTGGTCGCTTC 236
          ::::: ::::: ::::: ::::: :::::
      56  nAlaMetAspArgHisLeuPheLeuArgAsnIlePheTrpThrIleValL 73
          ::::: :::::
      237  CGGTATG..... 243
          ::::: :::::
      73  euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyraSPtyr 89
          ::::: ::::: ::::: ::::: :::::
      244  ....GGGGGACTGCCCGCAGGTTGGGGGCCGIGATTTTGTTGTGCTC 288
          ::::: ::::: ::::: ::::: :::::
      90  ILeleuGlyArgArg..... 94
          ::::: :::::
      289  ATAGTGGGCTCCATGGGGTCCGGCAAAATATGCTTGGCGGATGCCCTC 338
          ::::: :::::
      95  .....AlaL 96
          ::::: :::::
      339  TCTCAAGATGGCCGACCCCAATCGCTTTGCGGCAAGAAGACCTTCGGGTCC 388
          ::::: :::::
      96  euAspAlaLeuThrIlePro.....Ala 103
          ::::: :::::
      389  TGGACCAAGCTGACCGACCCCTCCGGGGGTCCGGCGCGTGTACACATCCAG 438
          ::::: :::::
      104  ValGly.....ProTyraSnArgTy 110
          ::::: :::::
      439  GCGGGCCCTACCGGACCCGTTCCAGACCCCGGAGCCCTCCCGATCACGGTTTA 488
          ::::: :::::
      110  rLeuThrArgValSerArgGlyCysAspValValGluLeuAsnProIleS 127
          ::::: ::::: ::::: :::::
      489  CTACGCCGTGTGGAGCGCGCTGCCGACGCGTGTCTTAACGACCGGT 538
          ::::: ::::: ::::: :::::

```

```

127 erAsnValaAspAspMetIleSerAlaAlaLysGluLysGluLysGly 143
    |||::: ::::: ::|||::: :::
539 CGGAGGCCCCCGAGATTGTCCGGGGGCTCCGAAGACGTCGGAAACAA 588
    |||::: ::::: ||||| :::
144 ProPheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspG 160
    |||::: ::::: ||||| :::
589 CCTTACAACCTGACCATCGCTTGGTTTCGATG.....GG 623
    |||::: ::::: ||||| :::
160 ygluAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGlyA 177
    |::: ||||| ||||| ||||| |||||:::
624 AGGCAACTGTGCTATCCCATCACGCGTCATGGAGTACACCGAATGCTCT 673
    |||::: |||::: ||||| ::::: |||
177 spvalGlnLeuLeuSerGluCysAlaValaGlnSerAlaGlnMetTrpAla 193
    ::: |||::: ||||| ::::: |||
674 ACAACAAGTCTCTGGGGGCTGTCCCATCCGAACGACGCCCGCTGG... 720
    |||::: ||||| ||||| ::::: |||
194 ValAspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLe 209
    ::||| ||||| ::: |||||::: |||:::
721 ..AACTACTATGACAGCTTCAGCGCCGTCAGCAGGATAACCTGGGTT 767
    |||::: |||::: ||||| ::: |||||
209 uThrIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrL 226
    :::: ::|||::: ::||| ||||| :
768 CCTGATGACGCGCCCGCGCTTTGAGACCGCGCGGACAGTACCTGGGCTCG 817
    |||::: |||::: ||||| |||||
226 euLysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
    ::|||::: |||::: ||||| |||||
818 TGAAGATAAACGACTGACGAGAGATTACACAGTTTATCTTGAGACACCGA 867
    |||::: ||||| ||||| |||||
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProse 257
    ::||| ||||| ::: |||::: |||||
868 GCCAAGGCGCTCTGT...AAGTACGCCCTCCCGCTCGGCATCCCCCGTTC 914
    |||::: ||||| ||||| |||||
257 rLysCysTrpThrThrGlnGlnTyrGlnThrGly..... 268
    ||||| ::: ||||| |||||
915 AGCCTGCTCTCCCCCAGCGCTACACAGAGGGGTGACGTTGACACAGA 964
    |||::: ||||| ||||| |||||
269 .....PheGlnGlyGlnHisLeuTyrProIleAlaAsp 279
    ||||| ||||| ||||| |||||
965 TCGGATGCTGCCCGCTCAATCCCCGAGAACACAGCGCACCGTCCGCTA 1014
    ||||| ||||| ||||| |||||
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
    ::: |||||
1015 TACAGCTTGAAGATCGCCGG..... 1035
    |||||
296 uGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
    ||||| ::|||::: |||||
1036 .....TGGCAC.....GGGCCCAAGGCCCA...T 1057
    |||||
313 roArgProAspSerValProGlnGluIleProAlaValThrLysLysAla 329
    ::||| |||||
1058 ACACGAGCACCTGTGTCGCCCGGAGCTG..... 1086
    |||||
330 GluGlyArgThrProAspAlaGluSerSerGluLysLysAlaProProG 346
    ::|||::: ||||| ::|||
1087 ...TCCGAGACCCCAACGCC....ACGCAAGCCAGAACTCGCCCCGGA 1127
    |||||
346 uAspSerGluAspAsp...MetGlnAlaGluAlaSerGluGluAsnProA 362
    ||||| |||||::: |||||
1128 AGACCCCGAGGATTGCGCCTCTTGGAGAGACCCCGTGGGACGTGGCGC 1177
    |||||
362 laAlaLeuProGluAspAspGluValProGluAspThrGlnHisAspAsp 378
    ::||| ::: |||||::: |||||
1178 CGCAATCCCAACCAACTGGCACATCCCGTGATCCAGAGCGCGCGACG 1227
    |||||
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProVa 395
    ||||| ||||| ||||| |||||
1228 CCTTACCATCCCCCGGCACCCCGAACAACATGGGCTGATC..... 1269
    |||||
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAla 412
    ::|||::: ||||| |||||
1270 .....GCCGGCGCGGTGGCGCGGCAGTCTCTCGGACG 1300
    |||||

```



```
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProVa 395
||| ||| |||:::|
1228 CCTTACCATCCCCGGCCACCCGACACATGGGCGCTGATC..... 1269
395 IGlulThrThrLysSerSerAsnAlaValSerMetProIlePheAla 412
|||:::| |||:::| |||:::| |||:::|
1270 .....GCCGGCGGGTGGCGGCAGTCTCCTGGCAG 1300
412 IapheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
|||:::| |||:::| |||:::| |||
1301 CCCTGGTCATTTC.....GGAATTGTCTACTGG 1329
```


OM of: US-09-994-064-11 to: N_Geneseq_032802:* out_format : pfs
Date: Sep 29, 2002 11:11 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+-p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09994064/runat_27092002_185414_13667/app-query.fasta_1.499
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdl -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09994064_@cgnl_1_622 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-994-064-11
Query length: 434
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 234.550000

score_list:

Sequence	Strd Orig	ZScore	EScore	len	! Documentation	..
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAQ68944	+	2311.00	4156.20	3.7e-223	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAQ33502	+	2311.00	4156.20	3.7e-223	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAQ81152	+	2311.00	4156.20	3.7e-223	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAQ67867	+	2311.00	4156.20	3.7e-223	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAQ49300	+	2311.00	4156.20	3.7e-223	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAQ33504	+	2311.00	4130.03	1.1e-221	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAQ44385	+	2311.00	4130.03	1.1e-221	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ22986	+	648.00	1135.44	6.8e-55	35	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:AAQ90524	+	200.00	333.42	3.2e-10	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAQ50489	+	198.00	329.86	5.0e-10	15	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:AAQ40070	+	193.50	321.62	1.4e-09	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAQ10959	+	190.00	318.71	2.1e-09	11	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:AAQ05677	+	190.00	316.37	2.8e-09	14	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ21883	+	188.50	308.91	7.4e-09	22	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAQ51318	+	182.50	302.12	1.8e-08	15	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAQ11074	+	169.00	279.59	3.2e-07	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ10209	+	166.50	275.07	5.7e-07	13	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ14455	+	166.50	266.04	1.8e-06	29	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:AAQ11073	+	165.50	274.06	6.4e-07	12	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:AAQ10194	+	165.00	273.24	7.2e-07	12	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAQ73488	+	162.50	269.55	1.1e-06	11	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ29115	+	162.50	269.52	1.2e-06	11	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAQ28191	+	162.50	261.35	3.3e-06	23	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAQ81210	+	162.50	250.37	1.3e-05	62	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAQ235418	+	162.50	250.37	1.3e-05	62	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:AAQ70993	+	159.00	258.99	4.4e-06	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:AAQ51325	+	159.00	258.99	4.4e-06	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAQ31670	+	158.00	260.27	3.8e-06	12	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAQ31670	+	158.00	257.18	5.6e-06	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAQ62131	-	158.00	237.79	6.7e-05	92	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAQ62164	+	158.00	231.68	0.0001	15	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:AAQ62175	+	158.00	231.05	0.0002	16	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:AAQ25519	+	158.00	206.16	0.0039	15	
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAQ10960	+	156.50	258.12	5.0e-06	11	

/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:AAQ50488 + 156.50 253.83 8.6e-06
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:AAQ25448 + 151.50 232.33 0.0001
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ42687 + 147.00 238.97 5.8e-05
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:AAQ48251 + 147.00 238.97 5.8e-05
/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:AAQ94354 + 147.00 238.97 5.8e-05

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:AAQ68944

seq_documentation_block:

ID AAQ68944 standard; DNA; 1305 BP.
AC AAQ68944;
XX
DT 13-APR-1995 (first entry)
XX
DE gD gene of infectious laryngotracheitis virus (ILT).
XX
KW Homology vector 586-36.6; gD gene; cassette gene; ss.
XX
OS Infectious laryngotracheitis virus.
XX
FH Key Location/Qualifiers
FT CDS 1..1305
FT /tag= a
XX
PN WO9419014-A.
XX
PD 01-SEP-1994.
XX
PF 28-FEB-1994; 94WO-US01826.
XX
PR 26-FEB-1993; 93US-0024156.
XX
PA (JAPG) NIPPON ZEON KK.
PA (SYTR) SYNTRO CORP.
XX
PI Cochran MD:
XX
DR WPI; 1994-294007/36.
DR P-PSDB; AAR58856.
XX
PT New recombinant fowl pox virus for use in vaccines - contains
PT genes expressing antigens of Newcastle disease virus and opt.
PT Infectious bronchitis virus
XX
PS Disclosure; Page 74-75; 85pp; English.
XX
CC AAQ68944 contains the coding region of the ILT gD gene and was
CC derived from an approx. 2060 bp EcoRI to BclI restriction sub-
CC fragment of the ILT KpnI genomic restriction fragment #8 (10.6kb).
CC It is part of a cassette in plasmid 586-36.6 constructed for
CC the purpose of inserting ILT gB and gD genes into the fowlpox
CC virus (FPV).
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

alignment_scores:

Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-994-064-11 x AAQ68944 ..

Align seg 1/1 to: AAQ68944 from: 1 to: 1305

1 MethisArgProHisLeuArgArgHisSerArgTyrTyrAlaIalysGlyGI 17
|||||
1 ATGCACTCTCTCATCTCAGACGGCACTCGCTACTACGGAAGAGA 50
|||||
17 uValleuAsnLysHisMetAspCysGlyGlyLysArgCysSergLya 34
|||||

51 GGTGCTTAACAACACATGGATTGCGTGGAACCGGTGCTGCTCAGCGC 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgIleHis 50
101 CAGCTGTATTCACTCTTTTCTGACCTGTGTACAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACGCTATGACCCGCCATTATTTTGGAGAAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACTATCGTACTGCTTCTCTCCGCTAGCCAGACACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGCCGCTGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrIleuThrArgValSerArgI 117
301 ATACCGCGGTTGGCCGCTATACAGATACCTCAGGTATCAAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTTGTGAGCTCAACCCGATTCTAACGTGACGACATGATAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGCGCGCCAAAGAAAGAGAAAGGGGGGCCCTTTCGAGGGCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
451 TGGTCTACGTGATTAAAGGCGACGACGCGGAGCAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
501 CTATAGAAAAGAGTACAGGAGATGTGGCGACGTACACTGCTATCTGAAT 550
184 ysaIaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCGCTCAATCTGCACAGATGTGGCGAGTGACTATGTCTTAGCACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
601 CTTGTATCGCGAATGGCGGCGGACTGACTATATCTCCCCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
651 GCTCTCTGCGCAATACTTCTGACCCCTGAAATCGGAGATTGGCGCAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTCGTAACCTAGAAAGTAAAGATCGGGTGC 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCCTCGAAATGCTGGACAACAGACAGTATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAGGCGCAACACCTTATCCGATCGCAGACACCAATACAGCAC 850
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGACGTATATCGGGGATAGCAAGATATTCTGACGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTGCTGAGGAAAAAGAAATCTAGCGCGCCAGACCCCTGCTCCAGATAG 950
317 rValProGlnIleIleProAlaValThrLysLysAlaGluGlyArgThrP 334
951 CGTCCCGCAAGAAATTCGCCGTGTAACCAAGAAAGCGAAGGCGCGACCC 1000

334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCGAGAAAGACAGCGAAAGAAAGAGGCCCTCCAGAAAGACTCGAGAGAC 1050
351 AspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAs 367
1051 GACATGCAGGACAGAGGCTTGTGAGAAATCTCGCCGCCCTCCCGAAGA 1100
367 pAspGluValProGluAspThrGluHisAspAspProAsnSerAspPro 384
1101 CGACGAGTCCCGAGAGACACGACGACGATGATCCAAACTCGGATCCTG 1150
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGCCCGCGTGATCCCGGTGGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGTCTCCATGCCCATATTGCGCGGCTTCGTAGCCTGCCG 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTCCGCGCTCGTGGGCTACTGTTGGAGCATCGTAAAAATGCGCGCGTA 1300
434 er 434
1301 GC 1302
seq_name: /SIDSl/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: AAT33505
seq_documentation_block:
ID AAT33505 standard; DNA; 1305 BP.
XX
AC AAT33505;
DT 19-NOV-1996 (first entry)
XX
DE Infectious laryngotracheitis virus gd gene.
XX
KW Infectious laryngotracheitis virus; ILTV; herpesvirus;
KW attenuation; vector; vaccine; chicken; poultry; immunisation;
KW glycoprotein gd; ds.
XX
OS Infectious laryngotracheitis virus.
XX
PN WO9508622-A1.
XX
PD 30-MAR-1995.
XX
PF 16-SEP-1994; 94WO-US10628.
XX
PR 24-SEP-1993; 93US-0126597.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Cochran MD, Wild MA;
XX
DR WPI: 1995-139591/18.
DR P-PSDB; AAW00638.
XX
PT Recombinant attenuated infectious laryngotracheitis virus - for use
PT in vaccines to protect poultry from infection from the virus, also
PT methods of distinguishing between vaccinated and naturally infected
PT birds
XX
PS Example 1; Page 106-107; 177pp; English.
XX
CC The gd gene (AAT33505) spans bases 8462-9766 of the unique short
CC region (see also AAT33504) of infectious laryngotracheitis virus
CC (ILTV) genomic DNA. It codes for a glycoprotein (AAW00638) of
CC approx. 48,477 mol.wt. that is homologous to pseudorabies
CC virus 950 and to gd from herpes simplex virus-1, Marek's

CC disease virus, IPV and bovine herpesvirus-1.1. Monoclonal
CC antibodies raised to ILTV react specifically with gd from
CC ILTV and also with ILTV gd expressed in herpesvirus of
CC turkeys (HVT) virus vector. ILTV gd expressed in the HVT
CC vector is useful as a subunit vaccine.
xx

50 Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

alignment_scores:

Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x AAT33505 ..

Align seg 1/1 to: AAT33505 from: 1 to: 1305

1 MethisArgProHisLeuArgHisSerArgTyrTrpAlaIysGly 17
1 ATGCACCGTCCCTCATCTCAGACGCACTCGCGTTACTACGCGAAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyA 34
51 GGTCCTTAACAACACATGATGCGGTGGAACACGCTGCTCAGCGC 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
101 CAGCTGTATTCACCTCTTCTCGACTGTGTCAAGATATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACCGCTATGACCGCCATTATTTTGAGGAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACTATCGTACTGCTTCTTCTTCCTCGCTAGCCAGACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGCCGCTCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
301 ATACCGCGGTTGGCCCGTATACAGATACCTCAGGTATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTTGTGAGGCTCAACCCGATTCTTAACGTGACGACATGATAT 400
134 eraAlaAlaLysGluLysGluGlyGlyProPheGluAlaSerValVal 150
401 CGCGGCGCAAAAGAAAAGAGAGGGGGCCCTTCGAGGCCCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
451 TGGTTCTACGTGATTAAAGGCGCAGCAGCGCAGACAGACAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
501 CTATAGAAAAGAGTACAGGAGATGTGGCGACGTACAACCTGCTATCTGAAT 550
184 ysaAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCCGTTCAATCTGCACAGATGTGGCAGTGAGCTATGTTCCTAGCACC 600
201 leuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
601 CTTGTATCGCAAAATGGCGGAGCTGACTATATTCTCCCACTGCTGC 650
217 aleuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234

651 GCTCTGCGCCAATACTTGCTGACCCCTGAANAATCGGAGATTTGGCCA 700
234 hralaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTCGTAACCTTAGAAGTTAACGATCGCTGTTAAAGATCGGCTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCGTGAATGCTGGACACAGAACAGATATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAGGCGAACACACCTTTATCGCATCGCAGACACCAATACAGAC 850
284 lsaAlaAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGAGCAGCGTATATCGGAGATACGAAGATATTCTGACGCGCGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTTCGTGAGAAAAAGAAATCCTAGCGCGCCAGACCCCTCGTCCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrp 334
951 CGTCCCGCAAGAAATTCGCCGCTGTAAACCAAGAAAGCGGAAGGCGCACCC 1000
334 roAspAlaGluSerSerGlyLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCAAGAAAGCAGCGAAAGAGGCCCTCCAGAAAGACTCGAGGAC 1050
351 AspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAs 367
1051 GACATGCGAGCAGAGGCTTCTGTGAGAAATCCTCGCCGCTCCCGAAGA 1100
367 pasGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
1101 CGAGCAAGTCCCCGAGAGACCGAGCAGATGATCCAACCTCGGATCCTG 1150
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGCCCGCGCTGATCCCGGTGGAGGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTATAGCCGCTCTCCATGCCCCATATTGCGGCGTTGCTAGCCTGCGC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGCCTCGTGGGCTACTGTTTGAGACATCGTAAAAATGCGCGGCTA 1300
434 er 434
1301 GC 1302
seq_name: /SID51/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAx81152
seq_documentation_block:
ID AAX81152 standard; DNA; 1305 BP.
AC AAX81152:
DT 07-SEP-1999 (first entry)
DE Seq ID No: 19 of US5925358.
XX Fowlpox virus; FPV; recombinant; vaccine; immunisation; chicken; NDV;
KW Newcastle disease virus; NDV; Fowlpox; Infectious laryngotracheitis; ds.
XX Fowlpox virus.
OS US5925358-A.
PN 20-JUL-1999.
XX


```
XX 07-JUN-1995; 95US-0484575.
PF
XX
PR 07-JUN-1995; 95US-0484575.
PR 26-FEB-1993; 93US-0024156.
PR 28-FEB-1994; 94WO-US02252.
XX
PA (SYTR ) SYNTRO CORP.
XX
PI Cochran MD, Junker DE;
XX
DR WPI; 1999-418249/35.
DR P-PSDB; AAY21984.
XX
PT Fowlpox viruses, useful as vaccines for immunization of
PT chickens/turkeys against Fowlpox and Newcastle disease virus
XX
PS Disclosure; Columns 77-82; 108pp; English.
XX
CC The invention relates to a recombinant fowlpox virus (FPV) comprising
CC a foreign DNA inserted into a region of the fowlpox virus genome
CC corresponding to a 2.8 kb EcoRI fragment, capable of being expressed in a
CC host cell. The virus is used as a vaccine for immunising chickens against
CC Newcastle disease virus (NDV), Fowlpox, and Infectious Laryngotracheitis.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;
```

alignment_scores:

Quality:	2311.00	Length:	434
Ratio:	5.325	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:
US-09-994-064-11 x AAX81152 ..

Align seg 1/1 to: AAX81152 from: 1 to: 1305

```
1 MethisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGly1 17
1 ATGCACCGCTCCATCTCAGACGGCACCTCGCGTTACTACGCCGAAGGAGA 50
17 uValleuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
51 GGTGCTTAACAACAACATGGATTGCGGTGGAACGGTGCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheThrCysValArgIleMetArgGluHis 50
101 CAGCTGTATTCACCTCTTTCTGACTTGTCAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTGTACGCAACGCTATGACCGCCCATTTATTTTGAGGAATGC 200
67 aPheTrpThrIleValleuLeuSerSerPheAlaSerGlnSerThrAlaA 84
201 TTTTGGACTATCGTACTGCTTCTCTCGCTAGCCAGAGACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGCTACGACTACATTTTAGCCGCTCGCGCGCTCGACGGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG1 117
301 ATACCGGCGGTTGGCCCGTATACAGATACCTCAGGGTATCAAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
351 CTGCGACGTTGTCAGACTCAACCCGATTCTTAACGTGAGACGACATGAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGGCGGCCAAGAAAGAGAGAGGGGGCCCTTTCGAGGCGCTCCGTCGTC 450
```

```
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProI1 167
451 TGGTTTCTACGTGATTAAGGCGCAGACGCGGAGAGCAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
501 CTATAGAAAAGAGTACAGAGGAATGTGGCAGCGTACACTGCTATCTGAAT 550
184 ysaAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCCGTTCAATCTGCACAGATGTGGCAGCTGACTATGTTCTTAGCACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
601 CTGTATCGCGAATGGCGCGGAGCTGACTATATTTCTCCCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
651 GCTCTCTGGCCAATACTTGTCTGACCCCTGAATAATCGGGAAGATTGGCMAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTGTAACCTAGAAAGTTAACGATCGCTGTTTAAAGATCGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCCTCGAAATGCTGACACAACAGATATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAAAGCGCAACACCTTTATCCGATCGACACCAATACAGAC 850
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGAGACGCTATATCGGGGATACGAGATATCTGCAGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTGCTGAGGAAAGAAATCCTAGCGCGCCAGACCCCTGCTCCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrp 334
951 CGTCCCGCAAGAAATTCGCCGCTGTAAACCAAGAAAGCGGAAGGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCGAAGACAGCGAAGAAAGGCCCTCCAGAAAGACTCGGAGGAC 1050
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
1051 GACATGCAAGCAGAGGCTTCTGAGAAATCCTGCCGCCCTCCCGGAAGA 1100
367 paspGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
1101 CGACGAAGTCCCCGAGAGACACGAGCAGATGATCCAAACTCGGATCCTG 1150
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrLys 400
1151 ACTATTACAATGACATGCCCCCGCTGATCCCCGGTGAGGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGCTCTCCATGCCCATAATTGCGGCGTTCGTAAGCTTGC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGCGTCTGGGGTACTGTTTGGAGCATCGTAATAATGCGGCGGTA 1300
434 er 434
1301 GC 1302
```


seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: AAC67867

seq_documentation_block:

ID AAC67867 standard; DNA; 1305 BP.

XX AAC67867;

XX 02-MAR-2001 (first entry)

XX Recombinant fowlpox virus-related polynucleotide, SEQ ID NO: 19.

XX Fowlpox virus; FPV; antiviral; antibacterial; vaccine;

KW Newcastle's disease; Marek's disease; infectious laryngotracheitis; ds.

XX Unidentified.

XX US6136318-A.

XX 24-OCT-2000.

XX 07-JUN-1995; 95US-0486414.

XX 26-FEB-1993; 93US-0024156.

PR 28-FEB-1994; 94WO-US02252.

XX (JUNK/) JUNKER D E.

PA (COCH/) COCHRAN M D.

XX Cochran MD, Junker DE;

XX WPI; 2000-686071/67.

XX New recombinant fowlpox virus useful as vaccines contains foreign DNA

PT inserted into specific non-essential region of the genome -

XX Disclosure; Column 77-80; 56pp; English.

XX The present sequence is provided in a specification relating to a

CC recombinant fowlpox virus (FPV) that comprises a foreign DNA inserted

CC within a 3.5 kb EcoRI fragment of the FPV genomic DNA. The foreign DNA

CC can be expressed in host cells infected with FPV. The recombinant FPV

CC may be used in vaccines to protect animals (especially chickens) against

CC fowlpox and, depending on the source of the foreign DNA, other diseases,

CC particularly Newcastle's disease, Marek's disease or infectious

XX laryngotracheitis.

SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

alignment_scores:

US-09-994-064-11 x AAC67867 ..

Align seg 1/1 to: AAC67867 from: 1 to: 1305

1 MethISArgProHISLeuArgArgHISSerArgTYrTYrAlaIysGly1 17
1 ATGCACCGTCTCATCTCAGACGCGACTCGCGTTACTACGCGAAGAGAGA 50
17 uValLeuAsnLysHISMetAspCysGlyGlyLysArgCysCysSerGlyA 34
17 GGTGCTTAACAACACATGATTCGGGTGGAACGCTGCTCAGCGC 100
34 IaAlaValPheThrLeuPheTrpThrCysValArgIleMetArgIuHis 50
101 CAGCTGTATTCACTCTTTCTGGAAGTGTGTCAGAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHISLeuPheLeuArgAsnAl 67

151 ATCTGCTTTGTACGCAACGCTATGGACCGCCATTATTTTTGAGGAATGC 200
67 apHeTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACTATCGTACTGCTTCTTCTCCTTCCGTAGCCAGACACCGCCG 250
84 IaValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTACGCTACGACTACATTTAGCGCGTCCGCCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgI 117
301 ATACCGCGGTTGGCCCGCTATTAACAGATACCTCAGGTTATCAAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTTGTCAGCTCAACCCGATTCTAACGCTGACGACATGATAT 400
134 eRaAlaAlaLysGluLysGluLysGlyProPheGluAlaSerValVal 150
401 CGGCGGCCAAAGAAAAAGAGAGGGGGCCCTTTCGAGGCGCTCCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
451 TGGTTCTACGTGATTAAGGGCAGCAGCGCGAGACAGTACTGTCCAAT 500
167 eTYrArgLysGluTYrArgGluCysGlyAspValGluLeuLeuSerGlu 184
501 CTATAGAAAGAGTACAGGAGATGTGGCAGCGTACACTGCTATCTGAAT 550
184 ySaIaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCCGTTCAATCTGCACAGATGTGGCAGGTGACATGTTCTTAGCACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAla 217
601 CTTGTATCGCGAATGGCGGGGACTGACTATATCTCCCGCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGln 234
651 GCTCTCTGGCCAATACTTGTCTGACCCCTGAAAAATCGGGAGATTGCGCAA 700
234 hRaIaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTGTAACCTAGAGTTAAGATCGCTGTTAAAGATCGGGTGC 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGln 267
751 CAGCTTAACCTTTTACCGTCGAAATGCTGGACAACAGACATATCAGAC 800
267 rGlyPheGlnGlyLysIleLeuTyrProIleAlaAspThrAsnThrArg 284
801 TGGATTTCAGAGCGAAGCACTTTATCCGATCCGACAGACCAATACAGC 850
284 tSaIaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGACGATATCGGGGATACGAAGATATCTGACGCGCTGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAsp 317
901 AATTGCTGAGGAAAAAGAAATCTTAGCGCGCCAGACCTCTCCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThr 334
951 CGTCCCGCAAGAAATCCCGCTGTAAACCAAGAAAGGGAAGGCGCACCC 1000
334 roAspAlaGluSerSerGlyLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCAAGAAAGCAGCAAAAGAAAGGCCCTCCAGAAAGACTCGAGAG 1050
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGlu 367
1051 GACATGACAGCAGAGGCTTCTGGAGAAATCTCTGCCGCCCTCCCGAAGA 1100

367 paspgluValProgluAspThrGluHisAspAspProAsnSerAspProA 384
1101 CGACGAGTAGTCCCCGAGACACGACGATGATCCAAACTCGCATCCTG 1150
384 sptYrTyraSnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGCCCCCGCGTGATCCCGGTGAGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGTCTCCATGCCCATATTCCGGCGGTTCGTAGCCTGCCG 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGGCTGCTGGGGCTACTGTTTGAGCATCGTAAATGCCGCGGTA 1300
434 er 434
1301 GC 1302
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:AAZ49300
seq_documentation_block:
ID AAZ49300 standard; cDNA; 1305 BP.
AC AAZ49300;
XX
DT 14-MAR-2000 (first entry)
XX
DE ILTV glycoprotein D (gd) gene.
XX
KW Fowlpox virus; FPV; recombinant; antigenic protein; expression;
KW infectious laryngotracheitis virus; ILTV; Newcastle disease virus; NDV;
KW Marek's disease virus; cytokine; promoter; homologous recombination;
KW homology vector; multivalent; live vaccine; glycoprotein D; ds.
XX
OS Infectious laryngotracheitis virus.
XX
FH Key location/Qualifiers
FT CDS 1..1305
FT /*tag= a
FT /product= "glycoprotein D (gd)"
XX
PN US6001369-A.
XX
PD 14-DEC-1999.
XX
PF 07-JUN-1995; 95US-0477459.
XX
PR 26-FEB-1993; 93US-0024156.
PR 28-FEB-1994; 94MO-US02252.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Junker DE, Cochran MD;
XX
DR WPI; 2000-071638/06.
DR P-PSDB; AAY58184.
XX
PT Recombinant fowlpox virus useful as a vaccine for immunizing fowl
PT against Marek's disease, Newcastle disease, infectious
PT laryngotracheitis virus and/or fowlpox
XX
PS Claim 5; Columns 77-80; 56pp; English.
XX
CC The invention relates to a recombinant fowlpox virus (FPV)
CC comprising a foreign DNA inserted into a 4.2 kb EcoRI fragment
CC of the fowlpox virus genome. The foreign DNA is capable of being
CC expressed in a host cell into which the fowlpox virus has been
CC introduced and encodes an antigenic protein. The antigenic protein
CC which may be expressed includes infectious laryngotracheitis virus
CC (ILTV) glycoprotein B (gB) or glycoprotein D (gD, AAY58184), Newcastle

CC disease virus (NDV) haemagglutinin (HN, AAY58182) or fusion (F) protein
CC (AAY58183) and Marek's disease virus gB or gD. The foreign DNA may
CC alternatively encode a cytokine such as chicken myelomonocytic growth
CC factor (CMGF) or chicken interferon (cIFN). The foreign DNA in the
CC recombinant FPV is under the control of one or more synthetic pox
CC promoters, enabling control of strength and timing of heterologous
CC gene expression. The synthetic pox virus promoters that may be used are
CC based on promoters of the vaccinia virus and include early promoter 1
CC (EP1), late promoter 1 (LP1), EP2 and LP2 (AAZ49291-249294,
CC respectively). The recombinant FPV is generated via homologous
CC recombination between FPV DNA and a homology vector containing the
CC foreign DNA flanked by FPV sequences. The recombinant fowlpox viruses of
CC the invention are used as multivalent live vaccines for immunising fowl
CC against Marek's disease virus, NDV, ILTV and/or fowlpox virus. The
CC present sequence represents the ILTV glycoprotein D (gd) gene.
XX
SQ Sequence 1305 BP; 331 A; 349 C; 344 G; 281 T; 0 other;

alignment_scores: Length: 434
 Quality: 2311.00 Gaps: 0
 Ratio: 5.325
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x AAZ49300 ..
Align seg 1/1 to: AAZ49300 from: 1 to: 1305

1 MethisArgProHisLeuArgArgHisSerArgTyrTrpAlaLysGlyL 17
1 ATGCACCGTCTCTATCTCAGACGGCAGCTCGCGTTACTACGCCGAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyA 34
51 GGTGCTTAACAACAACATGATGCGGTGGAACACGGTGCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
101 CAGCTGTAATCACTCTTTCTTGACTTGTGAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACGCTATGACCGCATTTATTTTGAGGAATGC 200
67 apHeTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
201 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGACGCCGCG 250
84 laValThrTyraSptYrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGGCCGTCGCCGCTCGACGCCGTAACC 300
101 IleProAlaValGlyProTyraSnArgTyrLeuThrArgValSerArgGl 117
301 ATACCGGCGGTGGCCCGTATTAACAGATACCTCACTAGGGTATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
351 CTGCGACGTTGTCGAGCTCAACCCGATTTCTTAACGTGAGCAGCATGAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGGCGGCCAAGAAAGAGAAGGGGGCCCTTCGAGGCCCTCCGTCGTC 450
151 TrpPheTyraValIleLysGlyAspAspGlyGluAspLysTyrCysProil 167
451 TGGTCTACGTGATTAAGGGCGAGCAGCGGAGGACAAGTACTGTCAAT 500
167 eTyraGlyLysGluTyraArgGluCysGlyAspValGlnLeuLeuSerGluC 184
501 CTATAGAAAGAGTACAGGGAATGTGGCGAGCTACAACTGCTATCTGAAT 550

PA (SYTR) SYNPRO CORP.
XX
PI Cochran MD, Wild MA:
XX
DR WPI: 1995-139591/18.
DR P-PSDB: AAW00630, AAW00631, AAW00632, AAW00633, AAW00634, AAW00635,
XX AAW00636, AAW00637, AAW00638, AAW00639, AAW00640, AAW00641 W00642.
PT Recombinant attenuated infectious laryngotracheitis virus - for use
PT in vaccines to protect poultry from infection from the virus, also
PT methods of distinguishing between vaccinated and naturally infected
PT birds
XX
PS Example 1; Page 79-94; 177pp; English.
XX
CC The unique short region (AAT33504) of infectious laryngotracheitis
CC virus (ILTV) genomic DNA contains genes (see also AAT33505 and
CC AAT33510-13) that are associated with ILTV virulence. A deletion in
CC those genes, esp. the glycoprotein g6 gene, glycoprotein g1 gene,
CC thymidine kinase gene, US2 gene, UL47-like gene or the glycoprotein
CC g60 gene, will attenuate the ILTV. A gene for a foreign antigen may
CC be inserted into the US2, UL47-like, ORF4, g6, g60 or g1 gene to
CC produce a recombinant ILTV. Recombinant ILTV may be used as a
CC multivalent vaccine, esp. for use in poultry.
XX
SQ Sequence 13473 BP; 3390 A; 3582 C; 3547 G; 2951 T; 3 other;

alignment_scores: Quality: 2311.00 Length: 434
 Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x AAT33504 ..

Align seg 1/1 to: AAT33504 from: 1 to: 13473

1 MethisArgProHisLeuArgHisSerArgTyrAlaLysGly1 17
|||||
8462 ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACCGCAAGAGA 8511
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
|||||
8512 GGTGCTTAACAAACACATGATGGGTGGAAAAACGGTGCTGCTCAGGCG 8561
34 laAlaValPheThrLeuPheThrPheCysValArgIleMetArgGluHis 50
|||||
8562 CAGCTGATTCACCTCTTCTGCACTGTGTGAGGATTATCGGGAGCAT 8611
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
8612 ATCTGCTTTGTACGCAACGCTATGGACCGCCATTATTTTGAGGAATGC 8661
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
|||||
8662 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGACACCGCCG 8711
84 lavalThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
8712 CCGTCACGTACGACTACATTTTAAAGCGGTCGCGCGCTGACGCGCTAAC 8761
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
|||||
8762 ATACCGCGCGGTGGCCCGTATTAACAGATACCTCAGTAGGATCAAGAG 8811
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
|||||
8812 CTGGACGTTGTGAGCTCAACCCGATTCTTAACGTGAGCAGCATGATAT 8861
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
|||||
8862 CGGCGGCCAAAGAAAAAGAGAGGGGGGCCCTTTCGAGGCCCTCCGTCG 8911

151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
|||||
8912 TGGTTCTACGTGATTAAAGGGCAGCAGCGCGGAGACAGTACTGTCCAAT 8961
167 eTyrArgLysGluTyrArgGluCysGlyAspValGluLeuLeuSerGluC 184
|||||
8962 CTATAGAAAAGAGTACAGAGGAATGTGGCAGCTGACTATGTTCTGAAT 9011
184 ysAlaValGlnSerAlaGlnMetrrPalValAspTyrValProSerThr 200
|||||
9012 GCGCGTTCATCTGCACAGATGTGGCAGTGGACTATGTTCTTAGCACC 9061
201 leuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaA 217
|||||
9062 CTTGTATCCGGAATGGCGCGGAGTACTATATCTCCCACTGCTGC 9111
217 aleuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGln 234
|||||
9112 GCTCTTGCCCAATACTTGTGCTGACCCGTGAATCGGAGATTTGCGCAA 9161
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
9162 CAGCTCTGTAAGTCTAGAAAGTTAACGATCCGCTTTAAAGATCGGGTCG 9211
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
|||||
9212 CAGCTTAAGTCTTTTACCCTGCAATGCTGGACACACAGACAGATACAG 9261
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArg 284
|||||
9262 TGGATTTCAGAGCGAACAACCTTATCCGATCGCAGACACCAATACAGAC 9311
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
|||||
9312 ACGCGAGCAGCATATATCGGGGATACGAGATATTCTGACGCGCTGAAT 9361
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
|||||
9362 AATTGCTGAGGAAAAAGAAATCTAGCGGCCAGACCCCTCGTCCAGATAG 9411
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
|||||
9412 CGTCCCGCAAGAAATTCGCCGCTGTAACCAAGAAAGCGAAGGGCGCACCC 9461
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
|||||
9462 CGGACGCAAGAAAGCAGCGAAAGAAAGAGGCCCTCCAGAAAGACTCGAGGAC 9511
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
|||||
9512 GACATGACAGCAGAGGCTTCTGGAGAAATCTCTGCCGCCCTCCCGAAGA 9561
367 pAspGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
9562 CGACGAAGTCCCGAGAGCACCGAGCAGCATGATCCAAACTCGGATCTCTG 9611
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
|||||
9612 ACTATTACAATGACATGCCGCCGTGATCCCGGTGGAGAGACTACTATAA 9661
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
|||||
9662 AGTTCTAATGCCGCTCCATGCCCATATTGCGGCGCTTCGTAGCCTGCCG 9711
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
|||||
9712 GGTGCGCGCTCGTGGGCGTACTGGTTTGAGACATCGTAAATGCGCGGTA 9761
434 er 434
||
9762 GC 9763


```
9662 AGTTCTAATGCCGCTCCATGCCCATATTGCGGCGTTCGTAGCCTGCGC 9711
417 aValAlaIeuValGlyIeuValTrpSerIleValIysCysAlaArgS 434
|||||
9712 GGTGCGGCTCGTGGGCTACTGTTTGAGCATCGTAAATGCGCGCGTA 9761
434 er 434
||
9762 GC 9763

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:AAT44385
seq_documentation_block:
ID AAT44385 standard; DNA; 18912 BP.
XX
AC AAT44385;
XX
DT 02-JUN-1997 (first entry)
XX
DE Infectious laryngotracheitis virus unique short + flanking region.
XX
KW ILTV; vaccine; vector; attenuation; poultry;
KW avian infectious bronchitis virus; Newcastle disease virus;
KW infectious bursal disease virus of chickens;
KW Marek's disease virus; herpesvirus; ss.
XX
OS Infectious laryngotracheitis virus USDA strain 8302.
XX
FH Key Location/Qualifiers
FT repeat_region 1..2909
FT /tag= a
FT /note= "short repeat region"
FT 697..1533
FT /tag= b
FT /label= SRORF2
FT /note= "US10 gene"
FT complement (2916..3605)
FT /tag= c
FT /label= SRORF1
FT /note= "short repeat open reading frame 1"
FT 2910..16003
FT /tag= d
FT /note= "unique short region"
FT 2796..2891
FT /tag= e
FT /note= "polyA signal for ORF1"
FT complement (2916..3605)
FT /tag= f
FT /label= ORF1(RC)
FT /note= "unique-like short 2 (US2) gene"
FT 3642..3645
FT /tag= g
FT /note= "TATA signal for ORF1 and ORF2"
FT 3675..3678
FT /tag= h
FT /note= "TATA signal for ORF1"
FT 3677..3680
FT /tag= i
FT /note= "TATA signal for ORF2"
FT 3694..5124
FT /tag= j
FT /label= ORF2
FT /note= "protein kinase gene"
FT 5126..5131
FT /tag= k
FT /note= "polyA signal for ORF2"
FT 5173..5178
FT /tag= l
FT /note= "polyA signal for ORF2"
FT 5210..7081
FT /tag= m
FT /label= ORF3
FT /note= "unique long 47 (UL47)-like gene"
FT 7158..7161
FT /tag= n
FT /note= "TATA signal for ORF4"
FT 6748..7080
FT /tag= o
FT /label= ORF4
FT complement (6774..7154)
FT /tag= p
FT /label= ORF4(RC)
FT 7158..7161
FT /tag= q
FT /note= "TATA signal for ORF5"
FT 7245..8123
FT /tag= r
FT /label= ORF5
FT /note= "glycoprotein gg gene"
FT 7245..7322
FT /tag= s
FT 7323..8120
FT /tag= t
FT 8200..8205
FT /tag= u
FT /note= "polyA site for ORF5"
FT 8289..8294
FT /tag= v
FT /note= "polyA signal for ORF5"
FT 8333..11290
FT /tag= w
FT /label= ORF6
FT /note= "glycoprotein g60 gene"
FT 9623..10363
FT /tag= x
FT /note= "repeat region consists of approx. 23
FT repeats of 30-36 bp"
FT complement (9584..10462)
FT /tag= y
FT /label= ORF6(RC)
FT 11091..11094
FT /tag= z
FT /note= "potential TATA signal for ORF7"
FT 11098..12402
FT /tag= aa
FT /label= ORF7
FT 12455..12448
FT /tag= ab
FT /note= "TATA signal for ORF8"
FT 12510..13598
FT /tag= ac
FT /label= ORF8
FT /note= "glycoprotein gi gene"
FT 12510..12575
FT /tag= ad
FT 12576..13595
FT /tag= ae
FT complement (13253..13786)
FT /tag= af
FT /label= ORF8(RC)
FT 13705..13708
FT /tag= ag
FT /note= "TATA signal for ORF9"
FT 13792..15291
FT /tag= ah
FT /label= ORF9
FT /note= "glycoprotein gf gene"
FT 13792..13845
FT /tag= ai
FT 13846..15288
FT /tag= aj
FT 15116..15119
FT /tag= ak
FT /note= "TATA signal for ORF10"
FT 15298..16080
FT /tag= al
```

FT /label= ORF10
FT repeat_region 16004..18912
FT /*tag= am
FT CDS 16129..17013
FT /*tag= an
FT CDS complement (17380..18216)
FT /*tag= ao
XX W09629396-A1.
XX 26-SEP-1996.
XX
XX 21-MAR-1996; 96WO-US03916.
XX
XX 06-JUN-1995; 95US-0468190.
XX 23-MAR-1995; 95US-0410121.
XX
XX (SYTR) SYNTR0 CORP.
XX
XX Cochran MD, Wild MA;
XX
XX WPI; 1996-443172/44.
XX P-PSDB; AAW06782, AAW06783, AAW06784, AAW06785, AAW06786, AAW06787,
DR AAW06788, AAW06789, AAW06790, AAW06791, AAW06792, AAW06793, AAW01415,
DR AAW01416.
XX
XX Recombinant infectious laryngotracheitis virus with deletion in the
PT glycoprotein G, g1 or US2 gene; etc. - useful for vaccines against
PT infectious laryngotracheitis in poultry
XX
XX Example 11; Page 138-154; 216pp; English.
PS
XX The nucleotide sequence of 19,912 bp of contiguous DNA (AAT44385)
CC from the unique short and flanking region of infectious
CC laryngotracheitis virus (ILTV) contains the entire 13,098 bp unique
CC short region (see also AAT44384) and includes 17 open reading frames
CC that encode proteins (AAW06782-94, AAW01415-16) of over 100 amino
CC acids, 10 of which show homology to other virus genes. Novel
CC recombinant, attenuated ILTV comprises the ILTV genome contg. a
CC deletion in the unique short region, esp. in the glycoprotein g1,
CC g1, US2, ORF4, UL47-like or g60 gene. The attenuated virus is
CC useful as a vaccine against ILTV. A foreign gene encoding an
CC antigen e.g. from another avian virus can be inserted into the US2,
CC UL47-like, ORF4, g6, g60 or g1 gene to provide a multivalent
CC vaccine for chickens and other poultry. Deletion of the g6 or g1
CC gene provides as a negative marker to distinguish vaccinated from
CC infected animals.
XX
XX Sequence 18912 BP; 4416 A; 5261 C; 5251 G; 3984 T; 0 other;
SQ

alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x AAT44385 ..

Align seg 1/1 to: AAT44385 from: 1 to: 18912

1 MethisArgProHisLeuArgAlaHisSerArgTyrTyrAlaIalysGlyG1 17
11098 ATGCACCGTCTCTCATCTCAGACGGCCTCGCTTACTACGCGAAGAGAGA 11147
17 uValLeuAnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
11148 GGTGCTTAACAACACATGATGCGGTGGAATAACGGTGTCTCAGGCG 11197
34 laAlaValPheThrLeuPheThrPheCysValArgIleMetArgGluHis 50
11198 CAGCTGATTCACCTCTTTCTGACCTTGTCTCAGGATTATGCGGAGCAT 11247

51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
11248 ATCTGCTTGTACCGACACGCTATGAGCCGCCATTTATTTTGAGGAATGC 11297
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
11298 TTTTGGACTATCGTACTGCTTCTTCCTCCGCTACCGACAGACCCGCCG 11347
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
11348 CCGTCACGTACGACTACATTTTAGCCCGTCGCCGCTCGACGCCGTAACC 11397
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG1 117
11398 ATACCGCGCGGTGGCCCGTAAACAGATACCTCAGGCTATCAAGAGG 11447
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
11448 CTGCGACGTTGTGAGCTCAACCCGATTCTTAACGTGAGACGACATGATAT 11497
134 erAlaAlaIysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
11498 CGCGCGCCAAAGAAAAAGAGAGGGGGCCCTTTCGAGGCCCTCCGTCGTC 11547
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI1 167
11548 TGGTTCTACGTGATTAAAGGCGACGACGCGGAGACAACTGCTCAAT 11597
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
11598 CTATAGAAAGAGTACAGGGAATGTGGCGACGTACCACTGCTATCTGAAT 11647
184 ysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
11648 GCGCCGTCATCTGCACAGATGTGGCGACGTAGCTATGTTCTTACGACCC 11697
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
11698 CTGTGATCGCGAAATGGCGGAGACTGACTATTTCTCCCCACTGCTGC 11747
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
11748 GCTCTGTGGCAATACTTGTGACCCGTGAATAACGGAGATTGGCCAAA 11797
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
11798 CAGCTCTCGTAACTCTAGAAGTTAACGATCGCTTTAAAGATCGGCTCG 11847
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
11848 CAGCTTAACCTTTTACCCTCGAAATGCTGGACACAGACAGATATCAGAC 11897
267 rGlyPheGlnGlyLysHisLeuTyrProIleAlaAspThrAsnThrArgH 284
11898 TGGATTTCAAAGGCGAACCTTTATCCGATCGCAGACACCAATACAGAC 11947
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTyrPsn 300
11948 ACGCGACGACGATATTCGGGATACGAAGATATTCTGACGCGCTGGAAT 11997
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
11998 AATTGCTGAGGAAAGAAATCTTAGCGCGCCAGACCCCTGTCAGATAG 12047
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
12048 CGTCCCGCAAGAATTCGCCGCTGAACCAAGAAAGCGAAGGCGCACCC 12097
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
12098 CGGACGCGAAGAAAGCAGCGAAGAAAGAGGCCCTCCAGAAAGACTCGAGAGAC 12147
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367

|||||
12148 GACATGCAGGAGAGGCTTCTGGAGAAATCCTGCCCTCCCGAAGA 12197
367 PASPGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
12198 CGACGAAGTCCCCGAGGACACCGAGCAGATGATCCAACTCGATCCTG 12247
384 sPTyrrTyrrAsnAspMetProAlaValIleProValGluGluThrLys 400
12248 ACTATTACAATGACATGCCCGCGCGGTGATCCCGGTGAGAGACTACTAAA 12297
401 SerSerAsnAlaValSerMetProIlePheAlaIlePheValAlaCysAl 417
12298 AGTTCTAATGCCGTCTCCATGCCCATATTCGGCGGCTTCGTAGCCTGCCG 12347
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
12348 GGTCCGCGCTCGTGGGGCTACTGTTGGAGCATCGTAAATGCCGCGCTA 12397
434 er 434
12398 GC 12399
seq_name: /SIDSl/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:AAQ22986
seq_documentation_block:
ID AAQ22986 standard; DNA: 3502 BP.
XX
AC AAQ22986;
XX
DT 02-NOV-1992 (first entry)
XX
DE Sequence of ILTV gp60 gene.
XX
KW Subunit vaccine; immunogen; glycoprotein; promoter; ss.
XX
OS Infectious laryngotracheitis virus.
XX
FH Key Location/Qualifiers
FT TATA_signal 294..299
FT CDS 339..3326
FT repeat_unit 1692..1712
FT repeat_unit /*tag= c 1755..1775
FT repeat_unit /*tag= d 1824..1841
FT repeat_unit /*tag= e 1857..1871
FT repeat_unit /*tag= f 1923..1943
FT repeat_unit /*tag= g 1986..2006
FT repeat_unit /*tag= h 2019..2039
FT repeat_unit /*tag= i 2115..2129
FT repeat_unit /*tag= j 2178..2198
FT repeat_unit /*tag= k 2211..2228
FT repeat_unit /*tag= l 2241..2261
FT repeat_unit /*tag= m 2304..2324
FT repeat_unit /*tag= n 2334..2351
FT repeat_unit /*tag= o

XX
PF 23-AUG-1991; 91WO-AU00383.
XX
PR 24-AUG-1990; 90AU-0001937.
XX
PA (WEBS-) WEBSTER A PTY LTD.
XX
PI Sheppard MG, Prideaux C, Johnson M, Fahey KJ, York JJ;
PI Kongsuwan K;
XX
DR WPI; 1992-096898/12.
DR P-PSDB; AAR22234.
XX
PT Vaccines against ILTV for use in chickens - comprises
PT non-infectious sub-units or live recombinant viruses and may be
PT administered by aerosol
XX
PS Claim 14; Fig 6; 122pp; English.
XX
CC The inventors claim a non-infectious subunit vaccine for use against
CC ILTV which comprises a glycoprotein of ILTV. They also claim DNA
CC encoding the vaccine, a synthetic polypeptide displaying the
CC antigenicity of the 205k complex or 60k ILTV glycoprotein and a
CC recombinant ILTV with heterologous DNA inserted into a non-essential
CC region of the genome; and a recombinant DNA mol. comprising an ILTV
CC promoter region operatively linked to a heterologous DNA sequence.
CC The promoter region is the ILTV gp60 promoter, 9p205 (9pB) promoter
CC or the ORF3 promoter region. KpnK/ORF3 is located 5' of the gp60
CC gene. The deduced AA sequence of the gp60 gene has 19 hydrophobic
CC AA residues at the N-terminus which may correspond to the signal
CC sequence. A second region of hydrophobic AAs (posn. 960 to 989) at
CC the C-terminus could function as a transmembrane anchor sequence.
CC There are nine potential N-linked glycosylation sites on the ILTV
CC gp60 protein. One of these sites at residue 677, may not be active
CC due to the presence of a proline residue within the N-X-S/T signal.
CC There are repeated sequences within the ILTV gp60 coding region.
CC They are not perfectly conserved but several copies had diverged so
CC that only 4 AAs out of 7 were conserved in all 13 copies of the
CC repeats. There is no significant homology between the protein and
CC any sequenced herpesvirus proteins in the database called PIR.
XX
SQ Sequence 3502 BP; 845 A; 969 C; 920 G; 768 T; 0 other;
alignment_scores:
Quality: 648.00 Length: 123
Ratio: 5.311 Gaps: 0
Percent Similarity: 99.187 Percent Identity: 96.748
alignment_block:
US-09-994-064-11 x AAQ22986 ..
Align seg 1/1 to: AAQ22986 from: 1 to: 3502
1 MethisArgProHisLeuArgArgHisSerArgTyrrAlaLysGlyG1 17
|||||
3134 ATGCACCGTCTCATCTCAGACGGCACTCGCGTTACTACCGGAAAGAGA 3183
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyA 34
|||||
3184 GGTGCTTAACGAACACATGATTCGGGTGGAAAAACGGTGCTGCTCAGCGG 3233
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
|||||
3234 CACGTGTATTACACTCTTTCTGGACTTGTGTCAAGATTATGCGGAGCAT 3283
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
3284 ATCTGCTTTGTACGCAACGCTATGAGACCGCATTTATTATTGAGGAATGC 3333
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
|||||
3334 TTTTGGACTATCGTACTGCTTCTTCTCCTTCGCTAGCAGACACCGCGG 3383

84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
3384 CCGTCACGTACGACTACATTTTAGCGCCGTCGCGCTCGACGCGCTAACCC 3433
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG1 117
|||||
3434 ATACCGCGCGGTGGCCCGTATACAGATACCTCAGTACGGTATCAAGAGG 3483
117 yCysAspValValGluLeu 123
|||||
3484 CTGCGAGCTTGTCTGAGCTC 3502

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: AAN90524

seq_documentation_block:
ID AAN90524 standard; DNA; 1602 BP.

XX AAN90524;
XX
DT 16-NOV-1989 (first entry)
XX
DE Herpes Simplex virus-1 gd gene and flanking regions.
XX
KW Herpes Simplex virus-1; gd gene; diagnostic test kit;
membrane-free polypeptide; antibody; anti-antibody; detection.
XX
OS Herpes simplex virus-1.
XX
FH Key Location/Qualifiers
FT CDS 237..1415
FT /*tag= a

PN AU8930061-A.
XX
PD 22-JUN-1989.
XX
PF 17-FEB-1989; 89AU-0030061.
XX
PR 30-AUG-1983; 83AU-0030061.
XX
PA (GETH) GENENTECH INC.
XX
PI Berman PW, Lasky LA;
XX
DR WPI; 1989-234099/33.
DR P-PSDB; AAP90165.
XX
PT Diagnostic test kit - contains molecularly cloned, truncated
membrane-free deriv. of polypeptide and complementary antibody
or anti-antibody.
XX
PS Disclosure; fig 1; 88pp; English.
XX
CC Herpes Simplex virus-1 gd gene (see corresp. AAP90887) and flanking
regions. This was used to make truncated membrane-free derivs.
CC with antigenic determinants capable of specifically binding
CC complementary antibody (Ab). The patent claims
CC a diagnostic kit contg. the deriv. and the Ab or anti-Ab,
CC for detecting HSV infection. See AAN90523-6, AAN90529-30, AAP90165,
CC and AAP90166.
XX
SQ Sequence 1602 BP; 310 A; 543 C; 462 G; 287 T; 0 other;

alignment_scores:
Quality: 200.00 Length: 489
Ratio: 0.943 Gaps: 21
Percent Similarity: 43.354 Percent Identity: 22.086

alignment_block:
US-09-994-064-11 x AAN90524 ..

Align seg 1/1 to: AAN90524 from: 1 to: 1602

2 HisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyGluVal 18
:::|||||
54 TACCGACACACGCGAGAACCC.....CCTAAGGGGAGG 88
18 IleuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyAla 35
:::|:::|
89 GCCATTTTACGAGGAGGAGGGGTATAACAAGTCTGTCTTAAAAAGCAG 138
35 laValPheThrLeuPheThrPheThrCysValArgIleMetArgGluHis... 50
:::|:::|
139 GGGTTAGGGAGTGTTCGGTCAAT.....AAGCTTCACGCGGACGACCA 182
51IleCysPheValArgAs 56
183 ACTACCCCGCATCATCAGTTATCTTAAGGTCTCTTGTGTGCGCTTC 232
56 naIaMetAspArgHisLeuPheLeuArgAsnaIaPheThrIleValL 73
:::|:::|
233 CGGTATG..... 239
73 euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
:::|:::|
240GGGGGGGCTGCCGCGCAGGTGGGGGCGGTGATTTGTTGTCGTC 284
90 IleLeuGlyArgArg..... 94
285 ATAGTGGCCTCCATGGGGTCCGCGGCAATATGCTTGGCGGATGCCCTC 334
95AlaL 96
335 TCTCAAGATGGCCGACCCCAATCGCTTTCGCGGCAAGACCTTCCGGTTT 384
96 euAspAlaLeuThrIlePro.....Ala 103
|||||
385 TGGACCGACGTGACCGACCTCCGGGGTCCGGCGGTGACACATCCAG 434
104 ValGlyProTyrAsnArgTyr.....Le 111
:::|:::|
435 GCGGGCCTACCGAACCCTGTCACGCCCCCAGCCTCCCGATCACGGTTTA 484
111 uThrArgValSerArgGlyCysAspValValGluLeuAsnProIleSera 128
|||||
485 CCGCCGTGTGAGCGCGCTGCCGCGCAGCGTCTCTTAACCGCACCGTCCG 534
128 snValAspAspMetIleSeraIaIaLysGluLysGlyGlyPro 144
:::|:::|
535 AGGCCCCCGATTTGTCGGGGGCTCCGAAAGACGTCCGGAACAACACC 584
145 pheGluAlaSerValValThrPheTyrValIleLysGlyAspAspGlyG1 161
:::|:::|
585 TACAACCTGACCATCGCTTGTTCGATG.....GGAGG 619
161 uAspLysTyrCysProIleTyrArgLysGluTyrArgLysGlyAspV 178
:::|:::|
620 CAACGTGTCTATCCCATCACGCTCATGGAGTACACCGAATGCTCTACA 669
178 alGlnLeuLeuSerGluCysAlaValGlnSeraIaGlnMetTyrAlaVal 194
:::|:::|
670 ACAAGTCTCTGGGGGCTGTCCCATCCGAACGACGCCCGCTGG..... 713
195 AspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLeuTh 210
:::|:::|
714 AACTACTATGACAGCTTCAGCGCGCTCAGCGAGGATTAACCTGGGGTTCT 763
210 rIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThLeuL 227
:::|:::|
764 GATGCACGCCCCCGCTTGTGAGACCGCGCGCAGTACCTCGGCTCGTGA 813
227 yslIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu.....Val 241
|||||
814 AGATAAAGCAGCTGGACGGAAGATTACACAGTTTATCTGAGACACCGAGCC 863

335 TCTCAAGATGGCGGACCCCAATCGCTTTCGGGCAAGACCTTCGGGTCC 384
96 euAspAlaLeuThrIlePro.....Ala 103
385 TGGACCAAGCTGACCGACCTCCGGGGGTCCGGCGGTGTACCATTCAG 434
104 ValGlyProTyrAsnArgTyr.....Le 111
435 GCGGGCTTACCGAACCCTTCCAGCCCCCAGCCTCCGATCACGGTTTA 484
111 uThrArgValSerArgGlyCysAspValValGluLeuAsnProIleSera 128
485 CCGCGGTGTGAGCGCGCTCCGCGAGCGTCTCTAAACGACCGCTCGG 534
128 snValAspMetIleSeraAlaAlaLysGluLysGluLysGlyPro 144
535 AGGCCCCCAGATGTCCGGCGGGGCTCCGAAGACGTCCGGAACAACCC 584
145 PheGluAlaSerValValTrpPheTyrValIleLysGlyAspGlyGly 161
585 TACAACCTGACCATCGCTTGGTTTCGGATG.....GGAGG 619
161 uAspLysTyrCysProIleTyrArgLysGluTyrArgLysGlyAspV 178
620 CAACTGTGCTATCCCATCAGGTCATGAGTACACCGAATGCTCCTACA 669
178 aGlnLeuLeuSerGluCysAlaValGlnSeraGlnMetTrpAlaVal 194
670 ACAAGTCTGTGGGGCTGTCCCATCCGAACGCAAGCCCGCTGG..... 713
195 AspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLeuTh 210
714 AACTACTATGACAGCTTCAGCGCCGTACGAGAGATAACTGGGGTCTCT 763
210 rIlePheSerProThrAlaAlaLeuSerGlyLysGluLeuThrLeuL 227
764 GATGCACGCCCCCGCTTGAGACCGCGGCACGTACCTGGCGCTCGTA 813
227 ysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu.....Val 241
814 AGATAACGACTGGACGGAGATTACACAGTTATCTTGAGACCGAGCC 863
242 AsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSerLy 258
864 AAGGCTCTCTGT...AAGTACACCTCCCGCTGCGCATCCCCCGTCAGC 910
258 sCysTrpThrThrGluGlnTyrGlnThrGly..... 268
911 CTGCCTCTCCCCCAGCCTTACAGCAGGGGTGACGGTGGACAGCATCG 960
269PheGlnGlyLysIleuThrPheProIleAlaAspThr 280
961 GGATGCTGCCCCGCTTCATCCCCGAGAACCAGCGACCGCTATATAC 1010
281 AsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLeuG 297
1011 AGCTTGAAGATCGCCGG..... 1028
297 nArgTrpAsnAsnLeuLeuArgLysLysAsnProSeraAlaProAspPro 314
1029 ...TGGCAC.....GGGCCCAAGGCCCA..TACA 1053
314 rgProAspSerValProGlnGluIleProAlaValThrLysLysAlaGlu 330
1054 CGAGCACCTGTGCCCCCGAGCTGCC..... 1082
331 GlyArgThrProAspAlaGluSerSerGluLysLysAlaProProGluAs 347
1083 ...GAGACCCCAAGCC.....ACGACGACGAACCTCGCCCGAAGA 1123
347 pSerGluAspAsp...MetGlnAlaGluAlaSerGlyLysAsnProAla 363

1124 CCCCAGGATTCGGCCCTTGTGAGGACCCCGTGGGAGCGGTGGCCGCC 1173
363 lAleuProGluAspAspGluValProGluAspThrGlnHisAspAspPro 379
1174 AAATCCCAACCAACTGGACATCCCGTGAATCAGAGACCGCGACGCCCT 1223
380 AsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProValG 396
1224 TACCATCCCCCGCCACCCCGAACAACATGGGCTGATC..... 1262
396 uGluThrThrLysSerSeraAlaValSerMetProIlePheAlaAlaP 413
1263GCCGGCGGCTGGGCGGAGTCTCTGCGAGGCC 1296
413 heValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
1297 TGGTCATTGTC.....GGAATTGTGTACTGG 1322
seq_name: /SIDSl/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:AAAN40070
seq_documentation_block:
ID AAAN40070 standard; DNA; 1608 BP.
XX
AC AAAN40070;
XX
DT 12-FEB-1992 (first entry)
XX
DE Sequence coding for Herpes simplex virus type 1 gd glycoprotein.
XX
KW Antigen; vaccine; immunogen; ss.
XX
OS Herpes simplex virus type 1.
XX
FH Key Location/Qualifiers
FT sig_peptide 241..315
FT mat_peptide 316..1425
FT /*tag= a
FT /*tag= b
XX
PN EP101655-A.
XX
PD 29-FEB-1984.
XX
PF 19-JUL-1983; 83EP-0401476.
XX
PR 06-JUL-1983; 83US-0510551.
PR 20-JUL-1982; 82US-0400028.
PR 25-OCT-1982; 82US-0436368.
PR 25-JAN-1984; 84US-0573642.
XX
PA (MOLE-) MOLECULAR GENETICS.
PA (AMCY) AMERICAN CYANAMID CO.
XX
PI Watson RJ, Wels JH, Enquist LW;
XX
DR WPI: 1984-057658/10.
DR P-PSDB; AAP40083.
XX
PT Antigenic herpes simplex polypeptide prodn. - by culturing
PT microorganisms contg. vector including viral DNA fragment
XX
PS Claim 23; Fig 3; 109pp; English.
XX
CC The inventors claim: (1) purified DNA sequences coding for the gd
CC glycoproteins; (2) recombinant DNA vectors contg. these Sqs;
CC (3) unicellular organisms contg. these DNA Sqs or vectors, and
CC (4) polypeptides prepd. by these transformed organisms. The prefd.
CC host organism is E.coli ATCC 39159, 39160, B-15451 or B-15471 (see
CC AAN40070/P40083) or B-15449 or B-15450 (see AAN40071/P40084).
XX
SQ Sequence 1608 BP; 313 A; 542 C; 465 G; 288 T; 0 other;

alignment_scores: Quality: 193.50 Length: 498
 Ratio: 0.908 Gaps: 21
Percent Similarity: 42.771 Percent Identity: 21.486

alignment_block:

US-09-994-064-11 x AAN40070 ..

Align seg 1/1 to: AAN40070 from: 1 to: 1608

```
2 HisArgProHisLeuArgHisSerArgTyrTyrAlaLysGlyGluVal 18
  :::::::::::::: :: :::::::::::::: ::::::::::::::
58 TACCGACACACCGACGAACC.....CCTAAGGGGAGGG 92
18 IleuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyAlaA 35
  :::: :::: :::: :::: :::: ::::
93 GCCATTTTACGAGAGAGGGGTATACAAAGTCTGCTTTAAAAAGCAG 142
35 IValPheThrLeuPheTrpThrCysValArgIleMetArgLuhis... 50
  ::::: ::::: ::::: ::::: ::::: :::::
143 GGGTTAGGAGTGTGTCGTCAT.....AAGCTTCAGCGCGAAGACCA 186
51 .....IleCysPheValArgAs 56
  ::::: ::::: ::::: ::::: ::::: :::::
187 ACTACCCCGATCATCAGTTATCCTTAAGGTCCTTTGTGTGTCGCTTC 236
56 nAlaMetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleVal 73
  ::::: ::::: ::::: ::::: ::::: :::::
237 CGGTATG..... 243
73 euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
  ::::: ::::: ::::: ::::: ::::: :::::
244 .....GGGGGACTGCCGCGAGGTGGGGCGCTGATTTGTTGTCGTC 288
90 IleuGlyArgArg..... 94
  ::::: ::::: ::::: ::::: ::::: :::::
289 ATAGTGGGCTCCATGGGGTCCGGCAATATGCCTTGGCGGATGCCCTC 338
95 .....AlaL 96
  ::::: ::::: ::::: ::::: ::::: :::::
339 TCTCAAGATGGCCGACCCCAATCGCTTTCGGCGCAAGACCTTCCGGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
  ::::: ::::: ::::: ::::: ::::: :::::
389 TGGACCAAGCTGACCGACCCCTCCGGGGGTCGGCGCGTACCACATCCAG 438
104 ValGly.....ProTyrAsnArgTyr 110
  ::::: ::::: ::::: ::::: ::::: :::::
439 GCGGGCCCTACCGGACCCGTTCCAGCCCCCAGCCCTCCGATCACGGTTTA 488
110 rleuThrArgValSerArgGlyCysAspValAlaGluLeuAsnProIles 127
  ::::: ::::: ::::: ::::: ::::: :::::
489 CTACGCCGTGTGGAGCGCCCTGCCGACGCGTGTCTTAACGACACCGT 538
127 erAsnValAspAspMetIleSerAlaAlaLysGluLysGlyLysGly 143
  ::::: ::::: ::::: ::::: ::::: :::::
539 CGGAGGCCCCCAGATGTGCGGGGGGCTCCGAAGACGTCCGGAAACAA 588
144 PropheGluAlaSerValValTrpPheTyrValIleLysGlyAspAsp 160
  ::::: ::::: ::::: ::::: ::::: :::::
589 CCTTACAACCTGACCATCGCTTGGTTTCGATG.....GG 623
160 yGluAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGly 177
  ::::: ::::: ::::: ::::: ::::: :::::
624 AGGCAACTGTGTATCCCATCACGGTCATGGAGTACACCGAATGCTCCT 673
177 spValGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTyrAla 193
  ::::: ::::: ::::: ::::: ::::: :::::
674 ACAACAAGTCTCTGGGGGCTGTCCCATCCGAACGACGCCCGCTGG... 720
194 ValAspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLe 209
  ::::: ::::: ::::: ::::: ::::: :::::
721 ...AACTACTATGACAGCTTCAGCGCGCTCAGCGAGGATTAACCTGGGTT 767
```

```
209 utrIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrL 226
  ::::: ::::: ::::: ::::: ::::: :::::
768 CCTGATGCACGCCCCCGCGCTTGTGAGACCGCCGCGACAGTACTCGCGCTCG 817
226 eulYsIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
  ::::: ::::: ::::: ::::: ::::: :::::
818 TGAAGTAAACGACTGACGAGGATTAACAGATTATCCCTGGAGCACCGCA 867
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProse 257
  ::::: ::::: ::::: ::::: ::::: :::::
868 GCCAAGGCTCCTGT...AAGTACGCCCTCCCGCTGCCGATCCCCCGCTC 914
257 rLysCysTrpThrThrGluGlnTyrGlnThrGly..... 268
  ::::: ::::: ::::: ::::: ::::: :::::
915 AGCCTGCCCTCTCCCCCAGGCTTACCAGAGGGGTGACGCTGGACAGCA 964
269 .....PheGlnGlyGluHisLeuTyrProIleAlaAsp 279
  ::::: ::::: ::::: ::::: ::::: :::::
965 TCGGATGCTGCCCGCGCTTATCCCGAAGAACGACGCGACCGTCCGCGTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
  ::::: ::::: ::::: ::::: ::::: :::::
1015 TACAGCTTGAAGATCGCCGG..... 1035
296 uGlnArgTyrPAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
  ::::: ::::: ::::: ::::: ::::: :::::
1036 .....TGGCAC.....GGGCCAAGGCCCA...T 1057
313 roArgProAspSerValProGlnGluIleProAlaValThrLysLysAla 329
  ::::: ::::: ::::: ::::: ::::: :::::
1058 ACACGAGCACCCCTGCTGCCCGCCGAGCTG..... 1086
330 GluGlyArgThrProAspAlaGlnSerSerGluLysLysAlaProProGl 346
  ::::: ::::: ::::: ::::: ::::: :::::
1087 ...TCCGAGACCCCAACGCC....ACGACGCCAGAACTGCCCGCGGA 1127
346 uAspSerGluAspAsp...MetGlnAlaGluAlaSerGlyGluAsnProA 362
  ::::: ::::: ::::: ::::: ::::: :::::
1128 AGACCCCGAGGATTCGGCCCTTGTGAGGACCCCGTGGGACGGTGGCGC 1177
362 lAlaLeuProGluAspAspGluValProGluAspThrGlnHisAspAsp 378
  ::::: ::::: ::::: ::::: ::::: :::::
1178 CGCAAAATCCACCAAACTGGCACATCCCGTCGATCCAGAGACGCCGCGAGC 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProVa 395
  ::::: ::::: ::::: ::::: ::::: :::::
1228 CCTTACCATTCCCCCGGCCACCCCGAACAACATGGCGCTGATC..... 1269
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAlaA 412
  ::::: ::::: ::::: ::::: ::::: :::::
1270 .....GCCGCGCGGGTGGGCGGCAGTCTCTGCGCAG 1300
412 lAheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
  ::::: ::::: ::::: ::::: ::::: :::::
1301 CCTGGTCAATTTC.....GGAATGTGTACTTG 1329
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: AAN60545

seq_documentation_block:

ID AAN60545 standard; DNA; 1608 BP.

XX

AC AAN60545;

XX

DT 02-JUL-1991 (first entry)

XX

DE Sequence of a portion of the Herpes Simplex Virus Type I (HSV-I)

DE

glycoprotein D (gD) gene.

XX

KW Probe; hybridisation; ds.

XX

OS Herpes simplex virus type I.

XX

FT	Key	Location/Qualifiers
FT	misc_feature	301..324
FT		/tag= a
FT		/label= probe A
FT	misc_feature	361..384
FT		/tag= b
FT		/label= probe B
FT	misc_feature	421..443
FT		/tag= c
FT		/label= probe C
FT	misc_feature	481..504
FT		/tag= d
FT		/label= probe D
FT	misc_feature	543..564
FT		/tag= e
FT		/label= probe E
FT	misc_feature	601..624
FT		/tag= f
FT		673..685
FT	misc_feature	/tag= g
FT		/label= probe G
FT	misc_feature	721..746
FT		/tag= h
FT		783..805
FT	misc_feature	/tag= i
FT		/label= probe I
FT	misc_feature	841..867
FT		/tag= j
FT		/label= probe J
FT	misc_feature	complement (181..206)
FT		/tag= k
FT		/label= probe R-1
FT	misc_feature	complement (397..417)
FT		/tag= l
FT		/label= probe A-1
FT	misc_feature	complement (458..480)
FT		/tag= m
FT		/label= probe B-1
FT	misc_feature	complement (517..540)
FT		/tag= n
FT		/label= probe C-1
FT	misc_feature	complement (576..600)
FT		/tag= o
FT		/label= probe D-1
FT	misc_feature	complement (634..658)
FT		/tag= p
FT		/label= probe E-1
FT	misc_feature	complement (696..718)
FT		/tag= q
FT		/label= probe F-1
FT	misc_feature	complement (721..746)
FT		/tag= r
FT		/label= probe G-3
FT	misc_feature	complement (757..780)
FT		/tag= s
FT		/label= probe G-1
FT	misc_feature	complement (815..840)
FT		/tag= t
FT		/label= probe H-1
FT	misc_feature	complement (883..900)
FT		/tag= u
FT		/label= probe I-1
FT	misc_feature	complement (916..963)
FT		/tag= v
FT		/label= probe J-3,J-1
FT	misc_feature	complement (994..1017)
FT		/tag= w
FT		/label= probe K-1
FT	misc_feature	complement (1059..1080)
FT		/tag= x
FT		/label= probe L-1
FT	misc_feature	complement (1118..1140)
FT		/tag= y

```

FT      misc_feature      /label= probe M-1
FT      complement (1175..1198)
FT      /*tag= z
FT      /label= probe N-1
FT      complement (1201..1223)
FT      /*tag= aa
FT      /label= probe O-1
FT      complement (1238..1260)
FT      /*tag= ab
FT      /label= probe P-1
FT      complement (1357..1380)
FT      /*tag= ac
FT      /label= probe Q-1
FT      complement (1411..1428)
FT      /*tag= ad
FT      /label= probe S-1
FT      complement (1513..1539)
FT      /*tag= ae
FT      /label= probe T-1

XX      WO8607387-A.
XX      PD      18-DEC-1986.
XX      PF      13-JUN-1986;      86WO-0001280.
XX      PR      13-JUN-1985;      85US-0744800.
XX      PA      (AMGE-) AMGEN.
XX      PA      (ABBO ) ABBOTT LABORATORIES.
XX      PA      (NOVO ) NOVO IND A/S.
XX      PI      Snitman DL, Stroupe SD;
XX      WPI; 1986-346610/52.
XX      DR      Nucleic acid hybridisation assays - suitable for assaying
PT      double-stranded DNA
XX      PS      Example; Table I, Pages 15-19; 47pp; English.
XX      CC      The examples illustrate the practice of the method of the invention.
XX      CC      Specifically demonstrated are hybridisation assays employing the two-
CC      probe system to detect and quantitate the amount of desired target
CC      sequence in a solution. Portions of the plus and minus strands of
CC      HSV-I gD gene were used as probes.
XX      SQ      Sequence 1608 BP; 313 A; 541 C; 466 G; 288 T; 0 other;

alignment_scores:
      Quality: 193.50      Length: 498
      Ratio: 0.908      Gaps: 21
Percent Similarity: 42.771      Percent Identity: 21.486

alignment_block:
US-09-994-064-11 x AAN60545 ..

Align seg 1/1 to: AAN60545 from: 1 to: 1608

      2 HisargProHisLeuArgargHisSerArgTyrAlaLysGlyGluVa 18
      ::::::::::|||
      58 TACCGACACACACCGACGACAC.....CCTAAGGGGAGGCG 92
      18 lleuasnlyshismetaspCysGlyGlyLysArgCysCysSerGlyAla 35
      93 GCCATTTTTACGAGGAGGAGGGGTATACCAAGTCTGCTTTAAAAAGCAG 142
      35 laValpethrleuphetrrpThrcysValArgIleMetargGluHis... 50
      ::||| ||||| ::::::|||:::
      143 GGGTTAGGAGTGTGTCGTCAT.....AAGCTTCAGCGCGAAGCAGCA 186
      51 .....IleCyspheValArgas 56

```

```
187 ACTACCCCGATCATCAGTTATCCTTAAGGCTCTTTTGTGTGGTGGCTTC 236
56 nAlaMetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleValL 73
237 CGGTATG..... 243
73 euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
244 .....GGGGGACTGCGCGCCAGGTGGGGGCGCGGTGATTTTGTTCGTC 288
90 IleLeuGlyArgArg..... 94
289 ATAGTGGGCTTCATGGGGTCCGGGCAATATGCTTGCGGATGCCCTC 338
95 .....AlaL 96
339 TCCTCAGATGGCCGACCCCAATCGCTTTCGGCGCAAGACCTTCCGGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
389 TGGACCAGCTGACCGACCTCCGGGGGTCCGGCGGTGTACCATCCAG 438
104 ValGly.....ProTyrAsnArgTy 110
439 GCGGGCTTACCGGACCGCTTCCAGCCCCCAGCCTCCCGATCACGGTTTA 488
110 rLeuThrArgValSerArgGlyCysAspValValGluLeuAsnProIles 127
489 CTACGCCGTGTGGAGCGCGCTCCCGCAGCGGTGCTCCTAAACGCACCGT 538
127 erAsnValAspAspMetIleSerAlaAlaLysGluLysGlyGly 143
539 CGGAGGCCCCCAGATTGTCCGGGGGCTCCGAAGACGTCCGGAAACAA 588
144 ProPheGluAlaSerValValTrpPheTyrValIleLysGlyAspAspGl 160
589 CCTACCAACCTGACCATCGCTTGGTTTCGATG.....GG 623
160 yGluAspLysTyrCysProIleTyrArgLysGluTyrArgGluCysGlyA 177
624 AGGCAACTGTGCTATCCCACTACGCTCATGGAGTACACCGAATGCTCCT 673
177 spValGlnLeuLeuSerGluCysAlaValGlnSerAlaGlnMetTrpAla 193
674 ACAACAAGTCTCTGGGGGCTCTCCCATCCGAACGCAGCCCCGCTGG... 720
194 ValAspTyrValProSer...ThrLeuValSerArgAsnGlyAlaGlyLe 209
721 ...AACTACTATGACAGCTTCAGCCCGCTCAGCGAGGATTAACCTGGGGTT 767
209 uThrIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrL 226
768 CTTGATGCACGCCCCCGCTTGGAGACCGCGGACGTACTGCGGCTCG 817
226 eulysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGlu..... 240
818 TGAAGATAAAGCACTGACGAGATTAACACAGTTTATCTGGAGACACCGA 867
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProse 257
868 GCCAAGGGCTCCTGT..AAGTACGCCCTCCCGCTGCGCATCCCCCGTTC 914
257 rLysCysTrpThrThrGluGlnTyrGlnThrGly..... 268
915 AGCCTGCTCTCCCCCAGGCTTACCAAGCAGGGGGTGACGGTGGACAGCA 964
269 .....PheGlnGlyGluHisLeuTyrProIleAlaAsp 279
965 TCGGGATGTCGCCCGCTTCATCCCCGAGAACCGCAGCGCTGCCGTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLe 296
```

```
1015 TACAGCTTGAAGATCGCCGG..... 1035
296 uGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
1036 .....TGGCAC.....GGGCCCCAAGGCCCA...T 1057
313 roArgProAspSerValProGlnGluIleProAlaValThrLysLysAla 329
1058 ACACGAGCACCTGCTGCCCCCGGAGCTG..... 1086
330 GluGlyArgThrProAspAlaGluSerSerGluLysLysAlaProProGl 346
1087 ..TCGAGACCCCAACGCC....ACGCAGCCAGAACTGCCCCCGGA 1127
346 uAspSerGluAspAsp..MetGlnAlaGluAlaSerGlyGluAsnProA 362
1128 AGACCCGAGGATTCGCCCTCTTGAGAGACCCCGTGGGAGCGTGGCCGC 1177
362 laAlaLeuProGluAspAspGluValProGluAspThrGluHisAspAsp 378
1178 CGCAATCCCAACCAACTGCGACATCCCGTCGATCGAGAGCGCCGAGC 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIleProVa 395
1228 CCTTACCATCCCCCGGCCACCCCGCAACATGGGCGCTGATC..... 1269
395 lGluGluThrThrLysSerSerAsnAlaValSerMetProIlePheAla 412
1270 .....GCCGGCGCGGTGGCGGCGCAGTCTCTGCGCAG 1300
412 lApheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
1301 CCTGTGTCATTGTC.....GGAATTGTGTACTGG 1329
seq_name: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:AAZ10959
seq_documentation_block:
ID AAZ10959 standard; DNA; 1185 BP.
XX
AC AAZ10959;
XX
DT 29-OCT-1999 (first entry)
XX
DE HSV-1 glycoprotein D coding sequence.
XX
KW Glycoprotein D; gd; recombinant antigen library; disease-related antigen;
KW multivalent antigenic polypeptide production; infection; allergen;
KW asthma; autoimmune disease; rheumatoid arthritis; diabetes; therapy;
KW multiple sclerosis; inflammatory condition; cancer; contraception;
KW immune response; ss.
XX
OS Herpes simplex virus.
XX
PN WO941383-A1.
XX
PD 19-AUG-1999.
XX
PF 10-FEB-1999; 99WO-US02944.
XX
PR 23-OCT-1998; 98US-0105509.
PR 11-FEB-1998; 98US-0021769.
PR 11-FEB-1998; 98US-0074294.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Bass SH, Howard R, Punnonen J, Stemmer WPC, Whalen RG;
DR WPI; 1999-518452/43.
XX
PT Recombinant multivalent antigenic polypeptide produced by
PT recombining nucleic acid sequences and screening, used in vaccines
PT against e.g. infections and cancer
XX
```

```
PS Example 7; Fig 7; 153pp; English.
XX
CC This sequence encodes the glycoprotein D (gD) from herpes simplex virus-1
CC (HSV-1). This sequence was used to create a recombinant antigen library.
CC The library comprises recombinant nucleic acids encoding antigenic
CC polypeptides and is produced by recombination of at least two forms of
CC nucleic acid, differing by at least two nucleotides, encoding a
CC disease-related antigenic polypeptide. The library can be used to produce
CC a recombinant multivalent antigenic polypeptides of the invention, that
CC contains at least two antigenic determinants (AD) from different
CC polypeptides. The multivalent antigenic polypeptides are used in vaccines
CC to induce a protective or therapeutic response to a wide variety of
CC infectious agents (bacteria, viruses, parasites, including Plasmodium
CC falciparum); allergens; asthma; autoimmune disease (e.g. rheumatoid
CC arthritis, diabetes, multiple sclerosis); other inflammatory conditions
CC and cancer, also, where directed against sperm antigens, they can be used
CC for contraception. The multivalent peptides can be evolved to induce an
CC optimised immune response against a wide variety of antigens,
CC particularly a broad spectrum response to many different strains of a
CC pathogen, including strains that are likely to appear in the future.
XX
SQ Sequence 1185 BP; 221 A; 431 C; 335 G; 198 T; 0 other;

alignment_scores:
    Quality: 190.00      Length: 336
    Ratio: 1.180        Gaps: 14
    Percent Similarity: 47.917    Percent Identity: 23.214

alignment_block:
US-09-994-064-11 x AAZ10959 ..

Align seg 1/1 to: AAZ10959 from: 1 to: 1185

106 ProTyrAsnArgTyrLeuThrArgValSerArgGlyCysAspValValG1 122
    ||| ::| ||| ::| ::|::|::|::|::|::|::|::|::|::|::|
235 CCGATCACGCGTTTACTACGCCCGTGTGGAGCGCGCTGCCGACGCGTGT 284
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
122 uLeuAsnProIleSerAsnValAspAspMetIleSerAlaAlaLysGluL 139
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
285 CCTAAACGACACCGTCGGAGGCCGCCAGATTGTCCGGCGGCGCTCCGAAG 334
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
139 ySGluLysGlyGlyProPheGluAlaSerValValTrpPheTyrValIle 155
    ::| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
335 ACGTCGGGAACAACACCTACACCTGACCATCGCTGTTCGGATG... 381
    ::| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
156 LysGlyAspAspGlyGluAspLysTyrCysProIleTyrArgLysGluTyr 172
    ||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
382 .....GGAGGCAACTGTGCTATCCCATCAAGGTCATGAGTA 419
    ||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
172 rArgGluCysGlyAspValGlnLeuLeuSerGluCysAlaValGlnSerA 189
    | |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
420 CACCGAATGCTCTACACAAGTCTGTGGGGCTGTCCCATCCGAACGC 469
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
189 laGlnMetTrpAlaValAspTyrValProSer...ThrLeuValSerArg 204
    ||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
470 AGCCCCGCTGG.....AACTACTATGACAGCTTCAGCGCCGTACGCGAG 513
    ||| ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
205 AsnGlyAlaGlyLeuThrIlePheSerProThrAlaAlaLeuSerGlyG1 221
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
514 GATAACCTGGGGTCTCGATGCACGCCGCCGCTTGGAGACCGCGGCAC 563
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
221 nTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnThrAlaLeuValT 238
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
564 GTACCTGCGGCTCGTGAAGATAAACGACTGAGCGAGATTACACAGTTTA 613
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
238 hrLeuGlu.....ValAsnAspArgCysLeuLysIleGlySerGlnLeu 252
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
614 TCCTGGAGCACCGACGCAAGGGCTCTCTGT...AAGTACGCCCTCCCGCTG 660
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
253 AsnPhleuProSerLysCysTrpThrThrGluGlnTyrGlnThrGly.. 268
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
661 CGCATCCCCCGTCAAGCTGCTCCCTCCGCCACGAGGCTTACACAGGGGGT 710
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
269 .....PheGlnGlyGluHisLeuT 275
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
711 GACGGTGACAGCATCGGAGTGTGCCCCGCTTCATGCCGAGAACACAGC 760
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
275 yTrProIleAlaAspThrAsnThrArgHisAlaAspAspValTyrArgGly 291
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
761 GCACCGTCGCCGTATACACTGAAGATCGCGGG..... 795
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
292 TyrGluAspIleLeuGlnArgTrpAsnAsnLeuLeuArgLysLysAsnPr 308
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
796 .....TGGCAC.....GGGCC 806
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
308 oSerAlaProAspProArgProAspSerValProGlnGluIleProAlaY 325
    |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
807 CAAGGCCCA...TACACGACACCCCTGTGCCCCCGAGCTG..... 846
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
325 alThrLysLysAlaGluGlyArgThrProAspAlaGluSerSerGlyLys 341
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
847 .....TCCGAGACCCCAAGCC.....ACGCAGCCA 873
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
342 LysAlaProProGluAspSerGluAspAsp...MetGlnAlaGluAlaSe 357
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
874 GAACTGCCCGCGAAGACCCCGAGATTCGCCCTCTGGAGGACCCCGT 923
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
357 rGlyGluAsnProAlaAlaLeuProGluAspAspGluValProGluAspT 374
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
924 GGGGACGGTGGCGCCGCAATCCACCAAACTGGACATCCCGTGCATCC 973
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
374 hrGluHisAspAspProAsnSerAspProAspTyrTrpAsnAspMetPro 390
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
974 AGGACGCGCGGACGCCCTTACCATCCCCGGCCACCCCGAACACATGGGC 1023
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
391 AlaValIleProValGluGluThrThrLysSerSerAsnAlaValSerMe 407
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1024 CTGATC.....GCCGCGCGGTGGCGG 1046
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
407 tProIlePheAlaAlaPheValAlaCysAlaValAlaLeuValGlyLeuL 424
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1047 CAGTCTCTGGCAGACCCCTGTCATTGTC.....GGAATTG 1081
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
424 euValTrp 426
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1082 TGTACTGG 1089
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

seq_name: /SIS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.AA005677
seq_documentation_block:
ID AA005677 standard; DNA; 1459 BP.
XX
AC AA005677;
XX
DT 07-JAN-1991 (first entry)
XX
DE Sequence encoding polypeptide with activity of surface protein gD.
XX
KW SS;
XX
OS HSV-1 Miyama strain.
XX
FH Key 186..1260
FT CDS /*tag= a
FT
XX
PN JP02182196-A.
XX
XX 16-JUL-1990.
XX
XX 15-DEC-1988; 88JP-0317546.
XX
XX 25-MAR-1988; 88JP-0072596.
XX
XX 19-JUL-1988; 88JP-0180114.
XX
XX 15-DEC-1988; 88JP-0317546.
```


XX (TAKE) TAKEDA CHEMICAL IND KK.
XX WPI; 1990-258264/34.
DR P-PSDB; AAR06489.

PT New polypeptide with similar biological activity - to surface
protein gp of HSV-1 Miyama strain produced by genetic technology
XX
PS Disclosure; Fig 1; 20pp; Japanese.

XX The sequence can be used to produce the recombinant protein in an
CC E.coli host.

XX
SQ Sequence 1459 BP; 275 A; 498 C; 423 G; 263 T; 0 other;

alignment_scores:
Quality: 190.00 Length: 336
Ratio: 1.180 Gaps: 14
Percent Similarity: 47.917 Percent Identity: 23.214

alignment_block:
US-09-994-064-11 x AAQ05677 ..

Align seg 1/1 to: AAQ05677 from: 1 to: 1459

106 ProTyrAsnArgTyrLeuThrArgValSerArgGlyCysAspValValG1 122
111 :::: ||| :::: :::: ||| :::: |||
345 CCGATCAGCGGTTACTACGCGCGTGTGGAGCGCGCTGCCGACGCGTGT 394
122 uLeuAsnProIleSerAsnValAspAspMetIleSerAlaAlaLysGluL 139
111 ||| :::: ||| :::: ||| :::: ||| :::: |||
395 CCAACGCGACCGTCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 444
139 ysgIuLysGlyGlyProPheGluAlaSerValValThrPheTyrValIle 155
445 ACGTCCGGAACAACCCCTACACCTGACCATCGCTGTTGCGGATG... 491
156 LysGlyAspAspGlyGluAspLysTyrCysProIleTyrArgLysGluTy 172
492 GGAGCAACTGTGCTATCCCATCACGCGCATGAGTA 529
172 rArgGluCysGlyAspValGlnLeuSerGluCysAlaValGlnSerA 189
111 ||| :::: ||| :::: ||| :::: ||| :::: |||
530 CACCGAATGCTCTACAACAAGTCTCTGGGGCTGTCCCATCCGAACGC 579
189 IagIleMetTrpAlaValAspTyrValProSer...ThrLeuValSerArg 204
580 AGCCCCCGCTGG.....AACTACTATGACAGCTTCAGCGCGCTCAGCGAG 623
205 AsnGlyAlaGlyLeuThrIlePheSerProThrAlaAlaLeuSerGlyG1 221
624 GATTAACCTGGGTTCTGATGACACGCGCGCGCTTTGAGACCGCGGCAC 673
221 nTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnThrAlaLeuValT 238
674 GTACCTGCGGCTCGTGAAGATAAACGACTGACGAGGAGATTACACAGTTTA 723
238 hrLeuGlu.....ValAsnAspArgCysLeuLysIleGlySerGlnLeu 252
724 TCCTGGAGCAGCAGCAGCGGCTCCTGT...AAGTACGCCCTCCCGCTG 770
253 AsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnThrGly.. 268
771 CGCATCCCCCGCTCAGCCCTGCTCTCCCCCAGGCGCTACACAGAGGGGT 820
269PheGlnGlyGluHisLeuT 275
821 GACGGTGGACAGCATCGGAGATGCTGCCCGCTTCATCCCGAGAACACAGC 870
275 yrProIleAlaAspThrAsnThrArgHisAlaAspAspValTyrArgGly 291
::: ||| :::: |||

871 GCACCGTCGCCGTATACAGCTTGAAGATCGCCGGG..... 905
292 TyrGluAspIleLeuGlnArgTrpAsnAsnLeuLeuArgLysLysAsnPr 308
906TGGCAC.....GGGCC 916
308 oSerAlaProAspProArgProAspSerValProGlnGluIleProAlaV 325
917 CAAGGCCCA...TACACGAGCACCCTGCTGCCCTGAGCTG..... 956
325 alThrLysLysAlaGluGlyArgThrProAspAlaGluSerSerGluLys 341
957TCCGAGACCCCAACGCC....ACGACGCCA 983
342 LysAlaProProGluAspSerGluAspAsp...MetGlnAlaGluAlaSe 357
984 GAACCTGCCCCCGGAGAGACCCCGAGAGATTGCGCCCTTGGAGAGACCCCGT 1033
357 rGlyGluAsnProAlaAlaLeuProGluAspAspGluValProGluAspT 374
1034 GGGAGCGGTGGCGCGCGCAATCCCACTGACATCCCGTCCGATCC 1083
374 hrGluHisAspAspProAsnSerAspProAspTyrTrpAsnAspMetPro 390
1084 AGGACGCCCGCAGCGCTTACCATTCGCCCGCCACCCCGAGACACATGGGC 1133
391 AlaValIleProValGluGluThrThrLysSerSerAsnAlaValSerMe 407
1134 CTGATC.....GCCGGCGCGGTGGCGG 1156
407 tProIlePheAlaAlaPheValAlaCysAlaValAlaLeuValGlyLeuL 424
1157 CAGTCTCTGGCAGCCCTGTCAATTTC.....GGAATTG 1191
424 euValTrp 426
1192 TGTACTGG 1199

.....

51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACGCTATAGACCGCCATTATTTTGGAGAAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
201 TTTTGGACTATCGTACTGCTTCTTCTTCTGCTAGCCAGACACCGCCG 250
84 IValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGCCGCTCGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
301 ATACCGGCGGTGGCCCGTATAACAGATACCTCACTAGGGTATCAAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTTGTGAGCTCAACCCGATTCTTACGTGAGACGACATGATAT 400
134 eAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGCGGCGCAAGAAAGAGAAAGGGGGCCCTTTCGAGGCGCTCCGTGTC 450
151 TrpPheTyrValIleLeuSGLYAspAspGlyGluAspLysTyrCysPro 167
451 TGGTTCTACGTATTAAGGGCGACGACGCGGAGCAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
501 CTATAGAAAGAGTACAGGGAATGTGGCGACGTACAACCTGCTATCTGA 550
184 yAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCGTCAATCTGCACAGATGTGGCAGTGGACTATGTTCTAGCACCC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAla 217
601 CTGTATCGCGAAATGGCGCGGACTGACTATATCTCCCCCACTGTGTC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGln 234
651 GCTCTGCGCAATCTGCTGACCCCTGAAATCGGGAGATTTGCCGAAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTGTAACCTAGAAAGTTAACGATCGCTGTTAAAGATCGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrGlnGlnTyrGlnThr 267
751 CAGCTTAACCTTTACCGTCGAATGCTGCACACAGACAGATATCAGAC 800
267 rGlyPheGlnGlyGlnHisLeuTyrProIleAlaAspThrAsnThrArg 284
801 TGGATTTCAAGGCGAACAACCTTATCCGATCGCAGACACCAATACAG 850
284 iAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGAGTATATCGGGGATACGAAGATATTCTGACGCGCTGGA 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAsp 317
901 AATTGCTGAGGAAAAAAGATCCTAGCGCGCAGACCCCTGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThr 334
951 CGTCCCGCAAGAAATCCCGCTGTAAACCAAGAAAGCGGAAGGCGCAC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCAAGAAAGCAGGAAAGAGCGCCCTCCAGAAGACTCGGAGAG 1050
351 AspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGlu 367

1051 GACATGACAGGAGGCTTCTGGAGAAATCCTGCCGCCCTCCCGAAGA 1100
367 pAspGluValProGluAspThrGlnHisAspAspProAsnSerAspPro 384
1101 CGACGAAGTCCCGAGAGACACCGAGACGATGATCCAACTCGATCCTG 1150
384 sPTyrTrpAsnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGATGCCCGCGCTGATCCCGGTGAGAGAGACTA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGTCTCCATGCCCATATTCGCGCGTTCGTAGCCTGCG 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArg 434
1251 GGTGCGCGCTCGTGGGCTACTGTTGGAGCAATCGTAAATGCGCGCTA 1300
434 er 434
1301 GC 1302
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-477-459-19
seq_documentation_block:
; Sequence 19, Application US/08477459
; Patent No. 6001369
; GENERAL INFORMATION:
; APPLICANT: Mark D. Cochran
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,459
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq. John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; US-08-477-459-19
alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x US-08-477-459-19 ..

Align seg 1/1 to: US-08-477-459-19 from: 1 to: 1305

```
1 MethIArgProHISLeuArgArgHisSerArgTyrTrpAlaLysGly 17
1 ATGCACCGTCTCATCTCAGACGGCAGCTCGGTTACTACGCCGAAAGGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGly 34
51 GGTGCTTAACAACACATGGATTGGGTGGAACGGTGCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheThrPThrCysValArgIleMetArgLuhIs 50
101 CAGCTGTATTCACCTCTTTCTGAGACTGTGTGAGGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTAGCAACGCTATGAGCCGCCATTATTTTGAGGAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACATATGCTACTGCTTCTCTCGCTAGCCAGACAGACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGCCGCTCGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArg 117
301 ATACCGCGGTGGCCCGTATACAGATACCTCCTAGGATCAAGAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTGTGAGCTCAACCCGATTTCTAACGTTGAGCAGCATGATAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGGCGGCCAAGAAAAGAGAAGGGGGCCCTTTCGAGGCCCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysPro 167
451 TGGTCTACGTGATTAAGGGCGACGCGCGGAGGACAACTGCTATCTGAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
501 CTATAGAAAGAGTACAGGAATGTGGCGAGCTGACCTATCTATCTGAAT 550
184 ySaLaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCCGTTCAATCTGCACAGATGTGGCGAGTGGACTATGTTCTAGCACC 600
201 leuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAla 217
601 CTGTATCGCGAATATGCGCGGAGCTGACTATATCTCCCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGln 234
651 GCTCTGTGCGCAATACTGCTGACCCCTGAATAATCGGGAGATTGCGCAAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTGTAAGTCTAGAGTTAAGTATGCTGTTAAAGATCGGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCGTGAATATGCTGACACACAGACAGTATCAGAC 800
267 rGlyPheGlnGlyLuhHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAGGCGAACACCTTATCCGATCGCAGACACCAATACACGAC 850
```

```
284 isAlaAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGACGACGTATATCGGGGATACGAAGATATCTGCAGCGCTGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTGCTGAGGAAAGAATCCCTAGCGCGCAGACCCCTCGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
951 CGTCCCGCAGAAATTCGCCGCTGTAACCAAGAAAGCGGAGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCGAAGACGACGGAAGAGGCCCTCCAGAAAGACTCGGAGGAC 1050
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
1051 GACATGACAGGACAGAGCTTCTGAGAAATCTGCGCCCTCCCGAAGA 1100
367 pasPglValProGluAspThrGluHisAspAspProAsnSerAspPro 384
1101 CGACGAAGTCCCGAGGACACCGAGCAGATGATCCAACCTCGATCCTG 1150
384 sPTyrrTyrAsnAspMetProAlaValIleProValGluGluThrTrlys 400
1151 ACTATTACAATGACATGCCCGCGTGATCCCGGTGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTCTAATGCCGCTCCATGCCCATATTCGGGGCTTCGTAGCCTGCGC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGCTCGTGGGCTACTGTTTGAGACATCGTAAATGCGCGCGTA 1300
434 er 434
1301 GC 1302

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-479-869-19
seq_documentation_block:
; Sequence 19, Application US/08479869
; Patent No. 6123949
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D, Mark D
; TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043.and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,869
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/024,156
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1305
US-08-479-869-19

alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-994-064-11 x US-08-479-869-19 ..

Align seg 1/1 to: US-08-479-869-19 from: 1 to: 1305

1 MethisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyL 17
1 ATGCACCGCTCCTCATCTCAGACGGCACCTGCGCTTACTACGCGAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyA 34
51 GGTGCTTAACAACAACATGATTGCGTGAAAAACGGTGCTGCTCAGCGG 100
34 laAlaValPheThrLeuPheThrCysValArgIleMetArgLysHis 50
101 CAGCTGTATTCACCTCTTTCTGACTGTGTCAAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACGCTATGACCGCCATTATTATTGAGGAATGC 200
67 aPheTTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACTATCGTACTGCTTCTTCTTCTGCTAGCCAGACACCGCGG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTACGACTACATTTTAGCGCGCTGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
301 ATACCGGCGGTTGGCCGCTAATACAGATACCTCAGTGGTATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
351 CTGCGACGTTGTGAGCTCAACCCGATTCTTAACGTGACGACACATGATAT 400
134 eraIaAlaLysGluLysGluLysGlyLysProPheGlnAlaSerValVal 150
401 CGGCGGCGCAAGAAAAAGAGAGGGGGCCCTTTCGAGGCGCTCGCTGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProI 167
451 TGGTCTTACGTGATTAGGGCGACGACGCGGCGAGACAACTGCTCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
501 CTATGAAAAAGACTACAGGGAATGTGGCGAGCTACCAACTGCTATCTGAAT 550
184 ysaIaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200

|||||
551 GCGCGCTTCATCTGCACAGATGTGGGAGTGACACTATGTTCTTACGACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaI 217
601 CTGTATCGCAAAATGGCGGGGAGTACTATATCTCCCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
651 GCTCTGTGGCCAATACTTGTGACCTGAAATCGGAGATTTGGCGAAA 700
234 hraIaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTCGTAACCTAGAAAGTTAACGATCGCTGTTAAGATCGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCCTGCAATGCTGACACACAGAACAGATACAGAC 800
267 rGlyPheGlnGlyLysHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAAAGCGCAACACCTTATCCGATCGCAGACACCAATACACGAC 850
284 isaIaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGACGATATATCGGGGATACGAAGATATTTCTGCAGCGCTGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTGCTGAGAAAAAGAAATCTTAGCGCGCCAGACCTCGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
951 GGTCCCGCAAGAAATTTCCCGCTGTAAACCAAGAAAGCGGAAGGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCAAGAAAGACGCAAAAGAGGCCCTCCAGAAAGACTCGGAGGAC 1050
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
1051 GACATGACGCGCAGAGGCTTCTGTGAGAAATCTCGCGCCCTCCCGAAAG 1100
367 pasGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
1101 CGACGAAGTCCCGCAGACACCGACGACGATGATCCAAACTCGGATCCCTG 1150
384 sPTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACATGACATGCCCGCGCTGATCCCGGTGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGTCTCCATGCCCATATTCGGCGGCTGTAGCCTGCGC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGCTCGTGGGCTACTGCTTGGAGCATCGTAAATGCGCGCGTA 1300
434 er 434
1301 GC 1302

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-486-414-19
seq_documentation_block:
; Sequence 19, Application US/08486414B
; Patent No. 6136318
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUSES AND USES THEREOF
; FILE REFERENCE: 42771D

; CURRENT APPLICATION NUMBER: US/08/486,414B
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Fowlpox virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(432)
US-08-486-414-19

alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-994-064-11 x US-08-486-414-19 ..

Align seg 1/1 to: US-08-486-414-19 from: 1 to: 1305

```
1 MethIsArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyL 17
|||||
1 ATGCACCGTCCCTCATCTCAGACGGCACCTCGCGTTACTACGCCGAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysSerGlyA 34
|||||
51 GGTGCTTAACAACACATGATGCGGTGGAAGCGGTGCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheThrCysValArgIleMetArgGluHis 50
|||||
101 CAGCTGTATTCACTCTTTCTGGACTTGTGTACAGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
151 ATCTGCTTTGTACGCAACGCTATGCAACGCCCATTTATTTGAGGAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAla 84
|||||
201 TTTTGGACTATCGTACTGCTTCTTCTTCCTGCTAGCCAGACACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgAlaLeuAspAlaLeuThr 100
|||||
251 CCGTCACGTACGACTACATTTTAGGCCGTCGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
|||||
301 ATACCGGCGGTGGCCCGTATACAGATACCTCAGGTAGGTTCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIle 134
|||||
351 CTGCGACGTGTGCGAGCTCAACCCGATTCTTAACGTGACGACATGATAT 400
134 erAlaAlaLysGluLysGluLysGlyLysProPheGluAlaSerValVal 150
|||||
401 CGGCGGCCCAAGAAAGAGAGAGGGGGCCCTTCGAGGCCCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProI 167
|||||
451 TGGTTCTACGTGATTAAAGGGCCAGCAGCGCGAGGACAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGlyCysGlyAspValGlnLeuSerGluC 184
|||||
501 CTATTAGAAAAGAGTACAGGGAATGTGGCGACGTACCAACTGCTATCTGAAT 550
184 ysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
|||||
551 GCGCCGTTCAATCTGCACAGATGTGGCGAGTGGACTATGTTCCTAGCACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
```

```
|||||
601 CTGTATCCGGAATATGGCGCGGAGTACTATATCTCCCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
|||||
651 GCTCTCTGCGCAATACTTGTGCTGACCCCTGAAGATCGGAGATTGGCGCAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
701 CAGCTCTCGTAACCTCTAGAGTTAAGATCGCTGTTAAAGATCGGGTGC 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
|||||
751 CAGCTTAACCTTTTACCCTGGAATCTGGACACACAGAACAGTATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
|||||
801 TGGATTTCAGAGCGAACACCTTATCCGATCGCAGACACCAATACACGAC 850
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
|||||
851 ACGCGGACGACGTATATCGGGGATACGAAGATATTCTGCACGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
|||||
901 AATTGCTGAGGAAAGAAATCTTAGCGCGCCAGACCCCTGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
|||||
951 CGTCCCGCAAGAAATTCGCCGTGTAACCAAGAAAGCGGAAGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
|||||
1001 CGGACGCAAGAACAGCAGCAAGAAAGAGCCCTCCAGAGACTCGGAGGAC 1050
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
|||||
1051 GACATGACGAGCAGAGGCTTCTGGAAGAAATCTTCCGCCCTCCCGAAGA 1100
367 pAspGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
1101 CGACGAGTCCCGGAGGACACCGACGATGATCCAAACTCGGATCCTG 1150
384 sPTrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
|||||
1151 ACTATTACAAATGACATGCCCGCGGTATCCCGGTGAGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
|||||
1201 AGTTCTAATGCCGTCTCCATGCCCATATTCCGCGGCTTCGTAGCCTGCC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
|||||
1251 GGTGCGGCTCGTGGGCTACTGTTTGAGAGCATCGTAAAGATCGCGCGTA 1300
434 er 434
1301 GC 1302
```

seq_name: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq: PCT-US94-01826A-19

seq_documentation_block:
; Sequence 19, Application PC/TUS9401826A
; GENERAL INFORMATION:
; APPLICANT: Syntro Corporation, et al.
; TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA

```

; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01826A
; FILING DATE: 28-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; PCT-US94-01826A-19
```

alignment_scores: Quality: 2311.00 Length: 434
 Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x PCT-US94-01826A-19 ..

Align seg 1/1 to: PCT-US94-01826A-19 from: 1 to: 1305

```

1 MethHisArgProHisLeuArgHisSerArgTyrTyrAlaLysGly1 17
1 ATGCACCGCTCCATCTCAGACGGCAGCTCGGTTACTACGCGAAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
51 GGTCCTTAACAACACATGGATTGGGTGGAACGCGTGCTGCTCAGGCG 100
34 laAlaValPheThrIleuPheTrpThrCysValArgIleMetArgLuhHis 50
101 CAGCTGTATTCACTCTTTTCTGGACTTGTGCAGGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTGTAGCAACGCTATGACCGCCCATTTATTTTGAGGAATGC 200
67 aPheTrpThrIleValIleuLeuSerSerPheAlaSerGlnSerThrAla 84
201 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGAGCACCGCCG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTAAGACTACATTTTAGCGCGTCGCGGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgG 117
301 ATACCGGGGTTGGCCCGTATACAGATACCTCAGTAGGGTATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspMetIleS 134
351 CTGCGACGTTGTGAGCTCAACCCGATTCTAACGTGAGACACATGATAT 400
```

```

134 eAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
401 CGGGCGCCCAAGAAAAAGAGAAGGGGGGCCCTTTCGAGGCCCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyLysAspLysTyrCysProIl 167
451 TGGTCTACGTGATTAGGGCGCAGCAGCGCGGAGGACAAGTACTGTCCAA 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
501 CTATAGAAAAGAGTACAGCGGAATGTGGCGACCTACAACCTGCTATCTGA 550
184 ysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCCGTTCAATCTGCACAGATGTGGGCAGTGGACTATGTTCTTAGCACC 600
201 leuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
601 CTTGTATCGCGAAATGGCGCGGAGCTGACTATATTCTCCGCCACTGCTGC 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
651 GCTCTCTGGCCAACTACTTGTCTGACCCCTGAATAATCGGAGATTTGCCCAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTCGTAACTCTAGAAGTTAACGATCGCTGTTAAAGATCGGGTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCGCTGAAATGCTGACACACAGACAGATATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAGGCGCAACACACTTTATCCGATCGCAGACACCAATACACGAC 850
284 isAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGACGTATATCGGGGATACGAAGATATTCGACGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTTGCTGAGGAAAAAGAAATCCTAGCGCGCCAGACCTCGTCCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
951 CGTCGCCGAAGAATTCGCCGTGTAAACCAAGAAAGCGGAAGGGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGCAAGAAAGCAGCGCAAGAAAGAGGCCCTCCAGAAAGACTCGGAG 1050
351 AspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAs 367
1051 GACATGCAGGACAGAGGCTTCTGGAGAAAATCCTGCCCGCCCGAAGA 1100
367 pasPgluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
1101 CGACGAAGTCCCGGAGGACACCGAGCAGATGATCCAAACTCGGATCCTG 1150
384 spTyrTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
1151 ACTATTACAATGACATGCCCGCGTATCCCGGTGGAGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
1201 AGTTCTAATGCCGCTCCATGCCCATATTTCGGCGGCTTCGTAGCCTGCC 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
1251 GGTGCGGCTCGTGGGGCTACTGTTGGAGCATCGTAAAAATGCGCGGTA 1300
434 er 434
```


11
1301 GC 1302

seq_name: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:PCT-US94-02252A-19

seq_documentation_block:
; Sequence 19, Application PC/TUS9402252A
; GENERAL INFORMATION:
; APPLICANT: Syntro Corporation, et al.
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/02252A
; FILING DATE: 28-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; PCT-US94-02252A-19

alignment_scores:
Quality: 2311.00 Length: 434
Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-994-064-11 x PCT-US94-02252A-19 ..

Align seg 1/1 to: PCT-US94-02252A-19 from: 1 to: 1305

1 MethisArgProHisLeuArgHisSerArgTyrTrpAlaIalysGlyG1 17
1 ATGCACCGTCCCTCATCTCAGACGCGACCTCGCGTTACTACGCCGAAGGAGA 50
17 uValleuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
17 GGTGCTTAACAACAACATGATGCGGTGGAAACGGTGTCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
101 CAGCTGTATTCACTCTTTCTGTGACTGTGTCAGGATTATCGCGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
151 ATCTGCTTTGTACGCAACGCTATGACCGCCCATTTATTTTGAGGAATGC 200

67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
201 TTTTGGACTATCGTACTGCTTCTTCCTTCGCTAGCCAGACGACCGCGC 250
84 laValThrTyraSPtyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
251 CCGTCACGTAAGACTACATTTTAGCGCGTCGCGCGCTCGACGCGCTAAC 300
101 IleProAlaValGlyProTyraSPtyrLeuThrArgValSerArgG1 117
301 ATACCGCGGTGGCCCGGTATACAGATACCTCAGTACGATGATCAAGAG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
351 CTGCGAGCTTGTGAGCTCAACCCGATTTCTAACGTGAGCAGCATGATAT 400
134 eAlaAlaLysGluLysGluLysGlyProPheGluAlaSerValVal 150
401 CGCGCGCCAAAGAAAAAGAGAAGGGGGCCCTTTCGAGCGCTCCGTCG 450
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
451 TGGTCTACGTGATTAAAGGCGACGCGCGAGAGACAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
501 CTATAGAAAGAGTACAGGGAATGTGCGACGCTACAACTGCTTCTGAAT 550
184 yAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
551 GCGCGCTTCAATCTGCACAGATGTGGCGAGTGACTATGTTCCTAGCAC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
601 CTTGTATCGCGAAATGGCGCGGACTGACTATATTTCTCCCGCACTGCT 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
651 GCTCTGTGGCCAATACTTGTGACCTGAAAAATCGGAGATTGTGGCAGAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
701 CAGCTCTCGTAACCTCTAGAAAGTAAAGATCGCTGTTAAAGATCGGGTC 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGlnGlnTyrGlnTh 267
751 CAGCTTAACCTTTTACCGTCGAATGCTGGACACAGAACAGATATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
801 TGGATTTCAGGCGGAACACCTTTATCCGATCGCAGACACCAATACAGAC 850
284 IsAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
851 ACGCGGACGACGTATATCGGGGATACGAAGATATTCTGCAGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
901 AATTGTGTGAGGAAAAAAGATCTTAGCGCGCAGACCCCTGTCAGATAG 950
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrP 334
951 CGTCCCGCAAGAAATTCCTGTAACCAAGAAAGCGGAAGCGCACCC 1000
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
1001 CGGACGAGAAAGACGAGCAAGAAAGAGCCCTCCAGAAAGACTCGGAGAG 1050
351 AspMetClnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAs 367
1051 GACATGACGAGCAGAGGCTTCTGAGAAAAATCTCTGCGCGCCCTCCGAGA 1100

```
367 pasgIuValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
1101 CGACGAAGTCCCGCAGACACCGACGACGATGATCCAAACTCGATCCTG 1150
384 sPTyRTyrAsnAspMetProAlaValIleProValGluGluThrTrlys 400
|||||
1151 ACTATTACAATGACATGCCCCGCGTGATCCCGGTGGAGAGACTACTAAA 1200
401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
|||||
1201 AGTCTAATGCCGTCTCCATGCCCATATTCGCCGCTTCGTAGCCTGCCG 1250
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
|||||
1251 GGTGGCGCTCGTGGGGCTACTGTTGGAGCATCGTAAATGCGCGCGTA 1300
434 er 434
||
1301 GC 1302
```

seq_name: /cgn2_6/plodata/2/ina/PCTUS_COMB.seq:PCT-US96-03916-10

```
seq_documentation_block:
; Sequence 10, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1305 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1305
; PCT-US96-03916-10
```

alignment_scores:
Quality: 2311.00
length: 434

Ratio: 5.325 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-994-064-11 x PCT-US96-03916-10 ..
Align seg 1/1 to: PCT-US96-03916-10 from: 1 to: 1305

```
1 MethIArgProHisLeuArgArgHisSerArgTyrTrpAlaLysGlyL 17
|||||
1 ATGCACCGCTCTCATCTCAGACGGCAGCTCGCGTTACTACGCGAAGAGA 50
17 uValLeuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
|||||
51 GGTGCTTAACAACAACATGGATTGCGGTGGAAGAAACGGTGCTGCTCAGGCG 100
34 laAlaValPheThrLeuPheTrpThrCysValArgIleMetArgGluHis 50
|||||
101 CAGCTGTATTCACCTCTTTCTTGACCTTGTCAGGATTATGCGGAGCAT 150
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
151 ATCTGCTTTGTACGCAACGCTATGACCGCCATTATTTTGAGGAATGC 200
67 aPheTrpThrIleValLeuLeuSerSerPheAlaSerGlnSerThrAlaA 84
|||||
201 TTTTGGACTATGCTACTGCTTCTTCTTCCTTCGCTAGCCAGACACCGCGG 250
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||
251 CCGTCACGCTACGACTACATTTTAGCGCCGTCGCGCGCTGACGCGCTAAC 300
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgI 117
|||||
301 ATACCGGCGGTGGCCCGTATAACAGATACCTCCTAGGGTATCAAGAGG 350
117 yCysAspValValGluLeuAsnProIleSerAsnValAspAspMetIleS 134
|||||
351 CTGCAGCGTTGTCAGACTCAACCCGATTTCTTAACGTGAGACGACATGAT 400
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVal 150
|||||
401 CGCGCGCCAAGAAAAGAGAAGGGGGCCCTTTCGAGGCCTCCGTCGTC 450
151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProIl 167
|||||
451 TGGTCTACGTGATTAAAGGCGCAGCAGCGCGGACGACAAGTACTGTCCAAT 500
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGluC 184
|||||
501 CTATAGAAAAGATACAGGAATGTGGCAGCTACAACGTCTATCTGAT 550
184 yAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
|||||
551 GCGCGGTTCAATCTGCACAGATGTGGCAGTGGACTATGTTCCTAGCACC 600
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl 217
|||||
601 CTGTATCGCGAATAGCGCGGAGCTGACTATATCTCCCCACAGCTGCG 650
217 aLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyArgPheAlaGlnT 234
|||||
651 GCTCTCGGCCAATACTGTGCTGACCCCTGAATACTGGGAGATTGGCGAAA 700
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
701 CAGCTCTCGTAACCTAGAAAGTTAACGATCGCTGTTTAAAGATCGGCTCG 750
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
|||||
751 CAGCTTAACCTTTTACCGTGAATGCTGACCAACAGAACAGTATCAGAC 800
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
```

```

|||||
801 TGGATTTCAAGGCGAAGACACCTTTATCCGATCCGACAGACACCATACAGAC 850
284 tsalaspaspvaltyrargglytyrgluaspihleuglnargtrpasn 300
|||||
851 ACGCGGACGACGTATATCGGGGATACGAGATATTTCTGCAGCGCTGGAAT 900
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspse 317
|||||
901 AATTTGCTGAGGAAAAAGAAATCTTAGCGCGCCAGACCCCTGCTCCAGATAG 950
317 tvalproglngluileproalaValThrLysLysAlaglulgarqThrp 334
|||||
951 CGTCCCGCAAGAAATTCCTCGCTGTAACCAAGAAAGCGGAAGGCGCACCC 1000
334 roaspalagluserSergLulLysLysAlaProProgluaspsSergLuas 350
|||||
1001 CGGACGCGAGAAAGCAGCGAAAGAGGCCCCCTCCAGAGACTCGGAGAGAC 1050
351 AspmetGlnAlaglualaserGlygluasnProAlaAlaleuProgluas 367
|||||
1051 GACATGCGAGGCGAGAGGCTTCTGGAGAAATCCTGCCGCCCTCCCGAAGA 1100
367 paspgluValProgluasPthrGluHisaspaspProasnSeraspProa 384
|||||
1101 CGACGAAGTCCCCGAGGACACCGACGATGATCCAAACTCGGATCCTG 1150
384 sPTyrTYrasnAspmetProalaValIleProvalGluGluThrLys 400
|||||
1151 ACTATTACAATGACATGCCCCGCCGTGATCCCCGGTGAGAGAGACTACTAA 1200
401 SerSerAsnAlaValSermetProIlePheAlaIaphValAlaCysAl 417
|||||
1201 AGTCTAATGCCGCTCTCCATGCCCCCATATTCGCGCGTTCGTAGCCTGCGC 1250
417 avAlAlaleuValGlyLeuLeuValTrpSerIleValLysCysAlaArg 434
|||||
1251 GGTGCGGCTCGTGGGCTACTGTGAGATCGTAATAATGCGGCGCTA 1300
434 er 434
||
1301 GC 1302

seq_name: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq: PCT-US96-03916-1
seq_documentation_block:
; Sequence 1, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:

```

```

: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 39116-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0525
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 13473 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: NO
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1059..2489
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2575..4107
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 4113..4445
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 4609..5487
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 5697..8654
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 9874..10962
:
: NAME/KEY: CDS
: LOCATION: 11159..12658
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 12665..13447
:
: PCT-US96-03916-1
:
alignment_scores:
    Quality: 2311.00      Length: 434
    Ratio: 5.325          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
alignment_block:
US-09-994-064-11 x PCT-US96-03916-1 ..
Align seg 1/1 to: PCT-US96-03916-1 from: 1 to: 13473
1 MethisArgProHisLeuArgArgHisSerArgTyrTyrAlaLysGlyL 17
|||||
8462 ATGCACCGTCCCTCATCTCAGACGGCAGCTCGCGTTACTACGGGAAGGAGA 8511
|||||
17 uValleuAsnLysHisMetAspCysGlyGlyLysArgCysCysSerGlyA 34
|||||
8512 GGTGCTTAACAACACATGATGGATTGGCGGTGGAACACGGTGCTGTCAGCGG 8561
|||||
34 laAlaValAlaPheThrLeuPheThrCysValArgIleMetArgLuhis 50
|||||
8562 CAGCTGTATTCACTCTTTTCTGGACTTGTGTCAGGATTATGCGGAGCAT 8611
|||||
51 IleCysPheValArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAl 67
|||||
8612 ATCTGCTTTGTACGCAACGCTATGAGACCGCCATTATTATTTGAGGAATGC 8661
|||||
67 aPheTrpThrIleValLeuLeuSerPheAlaSerGlnSerThrAlaA 84
|||||
8662 TTTTGGACTATCGTACTGCTTTCTTCCTTCGCTAGCCAGACACCGCCG 8711
|||||
84 laValThrTyrAspTyrIleLeuGlyArgArgAlaLeuAspAlaLeuThr 100
|||||

```

8712 CCGTCACGTACGACTACATTTTAGCCCGTCGGCGCTCGACGGCTAAC 8761
101 IleProAlaValGlyProTyrAsnArgTyrLeuThrArgValSerArgI 117
|||||
8762 ATACGGCGGTGGCCGTATTAACATACCTCACTAGGGTATCAAGAG 8811
117 YCysAspValAlaGluLeuAsnProIleSerAsnValAspMetIleS 134
|||||
8812 CTGCGACGTTGTGAGCTCAACCCGATTCTTACGTGACGACATGATAT 8861
134 erAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerVal 150
|||||
8862 CGGCGGCCAAAGAAAGAGAAGGGGGCCCTTCGAGGCCCTCCGTCGTC 8911
151 TrpPheTyrValIleLysGlyAspAspGlyLysPlystTyrCysProI 167
|||||
8912 TGGTCTACGTGATTAAGGGCGACGACGGCGGACAGCAAGTACTGTCCA 8961
167 eTyrArgLysGluTyrArgGluCysGlyAspValGlnLeuLeuSerGlu 184
|||||
8962 CTATGAAAAGAGTACAGGGAATGTGGCGACGTACAACCTGCTATCTGA 9011
184 ySaLaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerTh 200
|||||
9012 GCGCCGTTCAATCTGCACAGATGTGGCAGTGCATATGTTCTTAGCACC 9061
201 LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAla 217
|||||
9062 CTTGTATCGGAAATGGCGCGGACTGACTATATCTCCCCACTGCTGC 9111
217 aLeuSerGlyIntTyrLeuLeuThrLeuLysIleGlyArgPheAlaGln 234
|||||
9112 GCTCTGCGCAATACTTGCTGACCCCTGAATAATCGGGAGATTGCGCAAA 9161
234 hrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySer 250
|||||
9162 CAGCTCTCGTAACCTAGAAAGTTAACGATCCGTGTTAAAGATCGGCTCG 9211
251 GlnLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGlnTh 267
|||||
9212 CAGCTTAACCTTTTACCCTCGAATGCTGCAGACAACAGACAGTATCAGAC 9261
267 rGlyPheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThrArgH 284
|||||
9262 TGGATTTCAAAGCGGAACACCTTTATCCGATCCGACAGACCAATACGAC 9311
284 iSaLaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsn 300
|||||
9312 ACGCGGACGACGTATATCGGGATACGAGATATTTCTGCAGCGCTGGAAT 9361
301 AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe 317
|||||
9362 AATTGCTGAGGAAAGAAATCTTAGCGCGCCAGACCCCTCGTCAGATAG 9411
317 rValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThrp 334
|||||
9412 CGTCCGCAAGAAATTCGCCGTGTAAACAAGAAAGCGGAAGGCGCACCC 9461
334 roAspAlaGluSerSerGluLysLysAlaProProGluAspSerGluAsp 350
|||||
9462 CGGACGCAAGAAAGCAGGAAAGAGGCCCTCCAGAAAGACTCGGAGGAC 9511
351 AspMetGlnAlaGluAlaSerGlyLysAsnProAlaAlaLeuProGluAs 367
|||||
9512 GACATGCAAGCAGAGGCTTGTGAGAAATCTGCGGCCCTCCCGAAGA 9561
367 pAspGluValProGluAspThrGluHisAspAspProAsnSerAspProA 384
|||||
9562 CGACGAGTCCCGAGAGACACCGACGATGATCCAAACTCGGATCCTG 9611
384 sPTyTyrAsnAspMetProAlaValIleProValGluGluThrThrLys 400
|||||
9612 ACTATTACAATGACATGCCCGCGGTGATCCCGGTGAGAGACTACTAAA 9661

401 SerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaCysAl 417
|||||
9662 AGTTCTAATGCGCTCTCCATGCCCATATTCGCGCGTTCGTAGCCTGCGC 9711
417 aValAlaLeuValGlyLeuLeuValTrpSerIleValLysCysAlaArgS 434
|||||
9712 GGTCCGCGCTGCGTGGGCTACTGTTGGAGCATCGTAAAAATGCGCGCGTA 9761
434 er 434
||
9762 GC 9763
seq_name: /cgn2_6/ptodata/2/ina/PCrUS_COMB.seq:PCr-US96-03916-59
seq_documentation_block:
; Sequence 59, Application PC/rUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCr/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET INFORMATION:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18912 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 697..1533
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1900..2784)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2916..3605)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3694..5124


```

: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 5210..7081
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 7245..8123
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 8333..11290
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 11098..12402
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 12510..13598
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 13792..15291
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 15298..16080
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 16129..17013
: OTHER INFORMATION:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: complement (17380..18216)
: OTHER INFORMATION:
PCT-US96-03916-59

```

```
alignment_scores: *
    Quality: 2311.00      Length: 434
    Ratio: 5.325          Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000
```

```
alignment_block:
US-09-994-064-11 x PCT-US96-03916-59 ..
```

Align seg 1/1 to: PCT-US96-03916-59 from: 1 to: 18912

1	MethIaArgProHISLeuArgAaRhISSerAArgTYrTYrAlaLysGlyGI	17
11098	ATGCACCGCTCTCATCTCAGACGCGCACTCGCGTTACTACGCGGAAGGAGA	11147
17	uValLeuAsnLysHISmetAspCysGlyGlyLysArgCysSerGIYA	34
11148	GGTGCTTAACAACAACACATGGATTCGGTGGA AAAACGGTGCTGCTCAGCGC	11197
34	IaIaValPheThrLeuPheTrpThrCysValArgIleMetArgGIuHIS	50
11198	CAGCTGATTCACCTCTTTCTCGACTTGTCAGAGATTATGCGGGAGCAT	11247
51	IleCysPheValArgAsnAlaMetAspAaRhISLeuPheLeuArgAsnAl	67
11248	ATCTGCTTTGTAGCGCAACGCTATGACCGCCATTATTTTTGAGGAAATGC	11297
67	apHeTrpThrIleValLeuLeuSerSerPheAlaSerGIInSerThAlaA	84
11298	TTTTTGGACTATCGTACTGCTTTCTTCTTCCTTCGCTAGCCAGACGACCGCG	11347
84	IaValThrTYrAspTYrIleLeuGlyArgArgAlaLeuAspAlaLeuThr	100
11348	CCGTCACGTCAGACTACATTGTAAGCCGCTGCGCGCTGCAGCGCGCTAACCC	11397

101	ILPRALAVAlGLyPROTyRRASnArGTyRLeuThrArGVAlSerArG	117
11398	ATACCGCGGTGGCCCGGTATAACAGATACCTCAGGTATCAAGAG	11447
117	YCAspValValGLuLeuAsnProILeserAsnValAspAspMetILes	134
11448	CTGGACGTTGTGCAGCTCAACCCGATTTCTAAAGTGGACGACATGATAT	11497
134	erAlaAlaLysGLuLysGLuLysGLyProPheGLuAlaSerValVal	150
11498	CGCGCGCCAAAGAAAAAGAGAGGGGGCCCTTTCGAGGCCCTCCGTGTC	11547
151	TRPheTyrrValILeLysGLyAspAspGLuAspLysTyrrCysProIL	167
11548	TGGTTCACGTGATTAAGGGCGACGCGGAGGACAAGTACTGTCCAAT	11597
167	eTyrrArGLysGLuTyrrArGLuCySGLyAspValGLnLeuLeuSerGLu	184
11598	CTATAGAAAGAGACTACAGGGAATGTGGCGACCTACAACCTGCTATCTGAAT	11647
184	ysAlaValAlnSerAlaGLnMetTrpAlaValAspTyrrValProSerThr	200
11648	GCGCGCTTCAATCTGCACAGATGTGGGACAGTGACTATGTTCTAGCACC	11697
201	LeuValSerArgAsnGLyAlaGLyLeuThrILePheSerProThrAlaAl	217
11698	CTTGATTCGCGAAATGGCGGGACTGACTATATTCTCCCCACTGCTGC	11747
217	AlEuSerGLyGLnTyrrLeuLeuThrLeuLysILeGLyArGPheAlaGLnT	234
11748	GCTCTCGGCCAATACTTGTCTGACCCCTGAAAAATCGGGAGATTTCGCGCAA	11797
234	hrAlaLeuValThrLeuGLuValAsnAspArgCysLeuLysILeGLySer	250
11798	CAGCTCTCGTAACCTCTAGAAGTTAACGATCGCTGTTAAAGATCGGGTCG	11847
251	GLnLeuAsnPheluProSerLysCySTRPrThrThrgLGLnTyrrGLnTh	267
11848	CAGCTTAACCTTTTACCGTCGAATGTGTGCACACAGAACAGTATCAGAC	11897
267	rgLypheGLnGLuHisLeuTyrrProILeAlaAspThrAsnThrArgH	284
11898	TGGATTTCAAGCGCAACACCTTTATCCGATCGCAGACACCAATACACGAC	11947
284	IsAlaAspAspValTyrrArgGLyTyrrGLuAspILeLeuGLnArgTTrpAsn	300
11948	ACGCGACGACGTATATCGGGATACGAGATATTCTGCAGCGCTGGAAT	11997
301	AsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAspSe	317
11998	AATTTGCTGAGGAAAAAGAATCTAGCGGCCAGACCCCTCGTCAGATAG	12047
317	rValProGLnGLuILeProAlaValThrLysLysAlaGLuGLyArgThrP	334
12048	CGTCCCGCAAGAAATTTCCCGCTGTAAACCAAGAAAGCGGAAGGGCCACCC	12097
334	roAspAlaGLuSerSerGLuLysLysAlaProProGLuAspSerGLuAsp	350
12098	CGGACCGAGAAAGCAGCGAAAAAGAGGCCCTCCAGAAAGACTCGGAGAGAC	12147
351	AspMetGLnAlaGLuAlaSerGLyGLuAsnProAlaAlaLeuProGLuAs	367
12148	GACATGCAGGCAAGAGCTTCTGAGAAATCTTGCAGCCCTCCCGGAAGA	12197
367	paSPGLuValProGLuAspThrGLuHisAspAspProAsnSerAspProA	384
12198	CGACGAAGTCCCCGAGGACACCGACGACATGATCCAAACTCGGATCTCTG	12247
384	sPTyrrTyrrAsnAspMetProAlaValILeProValGLuGLuThrThrLys	400
12248	ACTATTACAATGACATGCCCGCCGTGATCCCGGTGGAGAGAGACTACTTAA	12297


```
241 ValAsnAspArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProse 257
      ::::: ||| ||| ::::: |||::: |||
868 GCCAAGGGCTCCTGT...AAGTACGCCCTCCGCTGCGCATCCCCCGTC 914
257 rLysCysTrpThrThrGlnGlnTyrGlnThrGly..... 268
      | ||| ::::: ||| ||| |||
915 AGCCTGCCCTCTCCCCCAGGCCCTACACAGAGGGGTGACGGTGACAGCA 964
269 .....PheGlnGlyGlnHisLeuTyrProIleAlaAsp 279
      ||| |||::: |||::: |||
965 TCGGATGCTGCCCGCTTCATCCCCGAGAACCAAGCCGCTCCCGCTA 1014
280 ThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIle 296
      ::::: |||
1015 TACAGCTTGAAGATCGCCGGG..... 1035
296 uGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAlaProAsp 313
      |||::: |||::: |||
1036 .....TGGCAC.....GGCCCAAGGCCCA...T 1057
313 roArgProAspSerValProGlnGlnIleProAlaValThrLysAla 329
      ::||| |||:::
1058 ACACGAGCACCTGTGTCGCCCGGAGCTG..... 1086
330 GlnGlyArgThrProAspAlaGlnSerSerGlnLysLysAlaProProG 346
      ::::: |||::: ||| ::::: |||
1087 ..TCCGAGACCCCAACGCC....ACGACGCCAAGACTCGCCCGCA 1127
346 uAspSerGluAspAsp..MetGlnAlaGlnAlaSerGlyGluAsnPro 362
      ||| ||| |||::: ||| ::::: |||
1128 AGACCCCGAGGATTCGGCCCTCTTGGAGAGACCCCGTGGGAGCGTGCG 1177
362 laAlaLeuProGluAspAspGlnValProGluAspThrGlnHisAspAsp 378
      ::::: ||| ::::: |||::: |||
1178 CGCAATCCCACTCAACTGCGACATCCGTCGATCCAGAGCGCCGCGACG 1227
379 ProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIlePro 395
      ||| ||| |||::: ||| ::::: |||
1228 CCTTACCATCCCCCGGCCACCCCGAACAACATGGCCCTGATC..... 1269
395 lGlnGlnThrThrLysSerSerAsnAlaValSerMetProIlePheAla 412
1270 .....GCCGGCGCGGTGGCGGCGAGTCTCTGCGCAG 1300
412 lApheValAlaCysAlaValAlaLeuValGlyLeuLeuValTrp 426
      ||::: ||| ||| |||::: |||
1301 CCCTGTCATTTC.....GGAATTGTGTACTGG 1329

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-793-958-3

seq_documentation_block:
; Sequence 3, Application US/08793958
; Patent No. 5814486
; GENERAL INFORMATION:
; APPLICANT: Cohen, Gary H.
; APPLICANT: Eisenberg, Roselyn J.
; APPLICANT: Nicola, Anthony
; TITLE OF INVENTION: Herpes Simplex Virus Glycoprotein D
; TITLE OF INVENTION: Variants
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/793, 958
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/499,568
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5814486and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32813
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 241..1422
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 316..1422
US-08-793-958-3

alignment_scores:
Quality: 193.50 Length: 498
Ratio: 0.908 Gaps: 21
Percent Similarity: 42.771 Percent Identity: 21.486

alignment_block:
US-09-994-064-11 x US-08-793-958-3 ..

Align seg 1/1 to: US-08-793-958-3 from: 1 to: 1608

2 HisArgProHisLeuArgArgHisSerArgTyrTrpAlaLysGlyGluVal 18
::: ||| ||| |||
58 TACCGACACACCGACGAACC.....CCTAAGGGGAGGCG 92
18 lLeuAsnLysHisMetAspCysGlyLysArgCysCysSerGlyAla 35
93 GCCATTTTACGAGAGGAGGAGGTATACAAAGTCTGTCTTTAAAGCAG 142
35 lValAlaPheThrLeuPheThrTrpThrCysValArgIleMetArgGlnHis... 50
::: ||| ||| |||
143 GGGTTAGGAGTGTTCGTCAT.....AAGCTTCAGCGGACGACCA 186
51 .....IleCysPheValArgAs 56
::: ||| ||| |||
187 ACTACCCGATCATCAGTATATCCTTAAGTCTCTTTGTGTGGTGGCTTC 236
56 nAlaMetAspArgHisLeuPheLeuArgAsnAlaPheThrThrIleVal 73
::: ||| ||| |||
237 CGGTATG..... 243
73 euleuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr 89
::: ||| |||::: |||::: |||
244 ....GGGGGAGCTGCCGCCAGGTTGGGGCCGTGATTTGTGTGTCGTC 288
90 lLeuGlyArgArg..... 94
|||::: ||| :::
289 ATAGTGGCCTTCATGGGTCCGCGCAATATGCTTGGCGGATGCCCTTC 338
95 .....AlaL 96
339 TCTCAAGATGGCCGACCCCAATCGCTTTCGGCGCAAGACCTTCGGGTCC 388
96 euAspAlaLeuThrIlePro.....Ala 103
```



```

306 CCTAAACGCACCCTCGGAGGGCCCCCAGATTGTCCGGGGGCCCTCCGAAG   355
139 ysglulysglylyprofhenglualaserValTPrPheTyrrValIle    155
      :::: |||::::: ::::::|||||| :::
356 ACGTCGGGAACAACCTTAACAACCTGACCATCGCTTGTTTCGGATG... 402
156 lysgluasparspciglualasplystyrcysProileTyrrArgLysgluty 172
      ||| :::: ||||| |||||
403 ..... GAGGCAACTGTGTATCCCATCACGCTCATGAGATA 440
172 rArgGlucYsgLYAsPVAlGlnLeuleuSerGIuCYSAlAValGINserA 189
      | ||||| ::: |||::: ||| :::::
441 CACCGAATGCTCTCAACAACAAGTCTTGGGGGGCTGTCCAATCCGAACGC 490
189 laGInMeTrPaLaVaLasPTyrrValProser..ThrlEuValSeraRg 204
      ||| ::||| ||| :: |||||:::
491 AGCCCCGCTGG.....AACTACTATGACAGCTTCAGCGCCGTCAAGCGAG 534
205 AsnGLyaLAglYLEuthrILEphEsErProTHraLaLaLeusErgLyGI 221
      ::::: |||::: :::: |||:::
535 GATAACCTGGGGTTCCTGATGCACGCCCCCgcgttgagaccggccggcac 584
221 nTYrrLEuLethrLEuLysILEglYArgPhEAglINTHraLaLeuVAIT 238
      ||||| :::|||||::: ::::: :::
585 GTACCTCGCGCTCGTAGAGATAAACGACTGACGAGAGATTACACAGTTTA 634
238 hrLEuglu....ValasnAPARgCYslEulsILEgLYSErgINleu 252
      ||||| ::::: ||| ||| ::: |||
635 TCCTGGAGCACCGAGCCAGAAGGCTCCTGT..AAGTACGCCCTCCCGCTG 681
253 AsnPheLUeUProSERlysCySTrpThrThRGluGlnTyrgLnThrgLy.. 268
      ::::: ||||| ||| ::: ::: ||||| |||
682 CGCATCCCCCGTCAGCCTGCCCTCTCCCCCAGGCGCTTACCAGCAGGGGGT 731
269 ..... PHeGlnGYlGHISLeUT 275
      ||| |||:::
732 GAACGGTGACAGCATCGGAGATGCTGCCCGCTTCATCCCCGAGAACCAAGC 781
275 yrPROILeaLASprTHRasnTHrarGHISAlaASpaSPvaLTyrARGLy 291
      ::||| ::: ::: |||
782 GCACCGTCGCCGTATACAGCTTGAAGATCGCCGGG..... 816
292 TYrGLuaSPIleUGlnaRGTPrasnSNLeULeuARGLysASNPR 308
      |||:::
817 ..... TGGCAC..... GGCGC 827
308 oSeRaLPraoSPProArGPraoSpsERvalProGInGUlleProALay 325
      |:::||||| ::||| |||:::
828 CAAGGCCCA..TACAGAGCACCCCTGTGCCCCCGAGAGCTG..... 867
325 alTHrLysLSalaglUgLYarGThrPROaspAlaglUsErSergLuLys 341
      ::::: |||||::: ||| :::::
868 ..... TTCGAGACCCCCCAAAGCC.... ACCGACCCA 894
342 LysAlaProPGIUASPseRgiUaSPAsp..MetGlnAlaglUALase 357
      ::: ||||| |||||::: ::: :::
895 GAACTCGCCCCGGAAGACCCCGAGGATTCGGCCCTCTTGAGAGACCCCGT 944
357 rgLYglUAsnPRoaLaLaLeuProGiUaSPAspgLUvAlProGiUaSpT 374
      ||| ::||| ::: ::::: |||:::
945 GGGGACGGTGGCGCCGAATCCCAACAAC TGACATACCGTCGATCC 994
374 hrcLIuHisAsPARPsPRoASnsERasRPraOSPtyRTyrASnaSPmetPro 390
      ::: ||| ||| |||::: |||
995 AGGACGCGCGACGCGCTTACCATTCCCCGGCCACCCCGAACAACATGGGC 1044
391 AlaValIlleProVALglUglUTHrThRLysSERasNaAlaValSerMe 407
      ::: :::: |||||:::
1045 CTGATC..... GCCGCGCGGTGGCGG 1067
407 tProIIepheaLaAlaphEyalaLaCYsaLaValaLaLeuValGIyleUL 424
      ::::: |||||::: ||| ||| :::::
1068 CAGTCTCTCTGGCAGCCCTGTGTCATTGTC.....GGAATTG 1102
```

```

424 euValTrp 426
      ::      |||
1103 TGTRACTGG 1110

seq_name: /cgn2_6/ptodata/2/ina/backfiles1.seq:5182195-9

seq_documentation_block:
; Patent No. 5182195
; APPLICANT: NAKAHAMA, KAZUO;KAISHO, YOSHIHIKO;YOSHIMURA, KOJI
; TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE
; DEFICIENT YEASTS
; NUMBER OF SEQUENCES: 71
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/269,140
; FILING DATE: 09-NOV-1988
; SEQ ID NO:9:
; LENGTH: 1459
5182195-9

```

```

alignment_scores:      quality: 190.00      Length: 328
                       Ratio: 1.173      Gaps: 14
Percent Similarity: 49.390      Percent Identity: 23.780

alignment_block:
US-09-994-064-11 x 5182195-9      ..

Align seg 1/1      to: 5182195-9      from: 1      to: 1459

106 ProTyrAsnArgTyrLeuThrArgValSerArgGlyCysAspValValG1 122
    ||      ::      |||      ::      :::::|||:::||||      |||
345 CCGATCAGCGGTTACTACGCCGTGTGGAGCGCGCTGCCGACGCTGTCT 394
    ||||||      |||:::      :::::      :::::      ::|||:::||||
395 CCTAACGCACCGTCGAGGCGCCCGACATGTCTCCGGGGGCTCCGAAG 444
    ||||||      |||:::      :::::      :::::      ::|||:::||||
122 uLeuAsnProIleSerAsnValAspAspMetIleSerAlaAlaIysGlu 139
    ||||||      |||:::      :::::      :::::      ::|||:::||||
395 CCTAACGCACCGTCGAGGCGCCCGACATGTCTCCGGGGGCTCCGAAG 444
    ||||||      |||:::      :::::      :::::      ::|||:::||||
139 ysgIuysgIglyPropheGluAlaSerValValATrphethyrValIle 155
    ::      |||:::      :::::      :::::      ::|||:::||||
445 ACGTCGGAAACAACCTTACAACCTGACCATCGCTGTGTTCCGATG... 491
    ::      |||:::      :::::      :::::      ::|||:::||||
156 LysgIyaspaspglygluAspIysTyrCysProIleTyrArgIysgluTy 172
    |||      ::      ::      ::      ||||||      |||||
492 .....GGAGCAACTGTGCTATCCCATCAGCGTCATGAGACTA 529
    |||      ::      ::      ::      ||||||      |||||
172 rArggluCysglyAspValGlnleuSerGluCysAlaValGlnSerA 189
    |      |||||:::      ::      |||:::      |||      :::::
530 CACCGAATGCTCTTACAACAAGTCTCTGGGGGCTGTCCCATCCGAAGC 579
    |||||      ::      ::      ::      |||||      |||||
189 laGlnMetTrpAlaValAspTyrValProSer...ThrIleuValSerArg 204
    |||      ::      ::      ::      ||||      :::::
580 AGCCCCGCTGG....AACTACTATGACAGCTTCAGCGCCGCTACGGAG 623
    |||      ::      ::      ::      ||||      :::::
205 AsnGlyAlaGlyLeuThrIlePheSerProThrAlaAlaIeuSerGlyG1 221
    :::::      |||:::      ::      :::::      ::|||:::||||
624 GATAACCTGGGGGTTCTCTGATGCACGCCCGCGCTTGAGACCGCGGCAC 673
    |||||      ::|||:::      ::|||:::      ::|||:::||||
221 nTyrIleuLeuThrIleuLysIleGlyArgPheAlaGlnThrAlaIeuValT 238
    |||||      ::|||:::      ::|||:::      ::|||:::||||
674 GTACCTCGGCTCGTGAAGATAAACGACTGGACGGAGATTACACAGTTTA 723
    |||||      ::|||:::      ::|||:::      ::|||:::||||
238 hrIleuGlu.....ValAsnAspArgCysLeuLysIleGlySerGlnIleu 252
    |||||      :::::      |||      |||      ::      |||
724 TCCTGGAGCACCGAGCCAGGCGCTCTCT...AAGTACGCCCTCCGCTG 770
    |||||      :::::      |||      |||      ::      |||
253 AsnPheIleuProSerLysCysTrpThrThrGluGlnTyrGlnThrGlyPh 269
    :::::      |||||      |||      ::      :::::      |||
771 CGCATCCCCCGTCAAGCCTGCTCTCCCGCCAGGCGCTTACCAGCAGGGGT 820
    :::::      |||||      |||      ::      :::::      |||
269 eGlnGlyGluHisLeuTyrProIle.....AlaAspThrAsnThrA 283
    ::      |||      ::      ::      ||||      ||||

```

821	GACGGTGGACAAGCTCGGGATGCTGCCCCGCTTCATCCCGGAACACG	870
283	rgnIsAlaAspAspValTtTargGlyTyrGluAspIleLeuGlnArgTtr	299
871	GC.....ACCGTCGCCGTATACAGCTTGAAGATCGCCGGGTGG	908
300	AsnAsnLeuLeuArgLysLysAsnProSerAlaProAspProArgProAs	316
909	CAC.....GGGCCCAAGGCCCA..TACACGAGCAC	937
316	pSerValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgT	333
938	CCTGCTGCCCCCTGAGCTG.....TCCGACA	963
333	hrProAspAlaGluSerSerGluLysLysAlaProProGluAspSerGlu	349
964	CCCCCAACGCC...ACGCAGCCAGAACTCGCCCGGAAGACCCCGAG	1007
350	AspAsp..MetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuPr	365
1008	GATTCGGCCCTCTTGAGAGACCCCGTGGGACGGTGGCGCGCAATGCC	1057
365	oGluAspAspGluValProGluAspThrGlnHisAspAspProAsnSerA	382
1058	ACCAAACTGGCACATCCCGTCGATCCAGGACGCCGCGACGCTTACCATC	1107
382	spProAspTyrTtTArgAsnAspMetProAlaValIleProValGluGluThr	398
1108	CCCCGGCCACCCCGAACAACAATGGCCCTGATC.....	1139
399	ThrLysSerSerAsnAlaValSerMetProIlePheAlaAlaPheValAl	415
1140GCCGGCGCGGTGGCGGACAGTCTCCTGGCAGCCCTGTTCAT	1180
415	acYsaAlaValAlaLeuValGlyLeuLeuValTtr	426
1181	TTCG.....GGAATTGTGTACTGG	1199

```
seq_name: /cgn2_6/prodata/2/1na/backfiles1.seq:5470718-1
```

seq_documentation_block:

```

; APPLICANT: O'CALLAGHAN, DENNIS J.
; TITLE OF INVENTION: EQUINE HERPESVIRUS TYPE 1 GLYCOPROTEIN
; D NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/954,417
; FILING DATE: 30-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 561,553
; FILING DATE: 01-AUG-1990
; SEQ ID NO:1:
; LENGTH: 2229
; 5470718-1

```

alignment_scores:		
Quality:	188.50	Length: 428
Ratio:	0.915	Gaps: 21
Percent Similarity:	48.131	Percent Identity: 23.131

```
alignment_block:
US-09-994-064-11 x 5470718-1 ..
```

Align seg 1/1 to: 5470718-1 from: 1 to: 2229

25 CysGlyGlyLysArgCysCysSerGlyAlaAlaValPheThrLeuPheTr 41
 ||| |||:::| | |::: ||| |||:::| |
 702 TGCCATGCGCATCGCGATCTTGAGCGTGTGCT.....CTCTGTG 742
 41 P-ThrCysValArgIleMetArgGluHisIleCysPheValArgAsnAla 57
 | |||||::: ||| |||||:::

743	GAACATGCGAGAAAGCCAAAGCGTGG.....	GTTCGAGGACGC	780
58	MetAspArgHisLeuPheLeuArgAsnAlaPheTrpThrIleValLeu	74	
781	CAGGATAGG.....	789	
74	userSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyr...I	90	
790	CCAAGGAGGTTCCACCACCCCGCTATACCTATACAA	826
90	IleLeuGlyArgArgAlaLeuAspAlaLeuThrIleProAlaValGlyPro	106	
827	TTTTAACAAAGATCAACGCGACTGCCCTAGTATCACCCTTCATTAAAGAC	876	
107	TyrAsnArgTyrIleuThrArgValSerArgGlyCysAs	119
877	CAAGTAAAAAATGTTGACTTGGCGATTGTTACTGCTACGCGCCCATGTGA	926	
119	pValValGlnLeuAsnProIleSerAsnValAspAspMetIleSerAla	136	
927	AATGATAGCGCTGATCGCTAAGACAACATAGACTCAATCCTGAAGAGC	976	
136	IalysGlnLysGlnLysGlyGlyProPheGluAlaSerValValTrpPhe	152	
977	TGGCGCGTGGCCCAAAAACCTTATGCC.....	GCCAGACTCACCTGGTTT	1020
153	TyrValIleLysGlyAspAspGlyGluAspLysTyrCys....	ProIl	167
1021	AAATTTATG.....	CCAACGTGTGCAACGCTAT	1049
167	eTyrArgLysGlnTyrArgGlnCysGlyAspValGlnLeuLeuSerGlnC	184	
1050	ACACGATGTTAGTTATATGAATGCAACCCGAAGCTATCATTTGCCATGT	1099	
184	ysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr	200	
1100	GTCATGAGAGATCAGACATACTATGGCAAGCTAGTTTAATTACTATGGCT	1149	
201	LeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAl	217	
1150	GCTGAACACTGCATGACATGAACCTGGACTTGTACTGGCAGCCCTGCACATTTC	1199	
217	AlaSerGlnTyrLeuThrLeuThrLeuLysIle..	GlyArgPheAlaG	233
1200	TGCCTCGGGACTGTATCGCGGTGTATAGAATCGACGGAAGCGCAATTT	1249	
233	InThrAlaLeuValThrLeuGlnValAsnAspArgCysLeuLysIleGly	249	
1250	ACACGGACTTTCCTGTACTATATCCACAGTGAACGGGTGT..	CCGATTGCC	1296
250	SerGlnLeuAsnPhe..LeuProSerLysCysTrpThrThrGlnGlnTyr	265	
1297	TTTGAGCTAAACTTTGGCAATCCGGATCGGTGTAAAACTCCAGACAGTA	1346	
265	rgInThrGly.....	PheGlnGlyGluHisLeuTyrP	276
1347	CTCGCGGGAGAGATTTTTACACGTGCGTTTCTTGTAATTCAACTTCC	1396	
276	roIleAlaAspThrAsnThrArgHisAlaAsp.....	Asp	287
1397	CACAAAGGAGAG.....	CATATGACATGGGTGAAGTTCTGCTTC	1434
288	ValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAsnAsnLeuLeuAr	304	
1435	GTCCTACGATGCTGGAACCTTCCACAGTCAGATTATGAAGCCAGCATTT	1484	
304	glySLysAsnProSerAlaProAspProArgPro....	AspSerValP	319
1485	CGCAAGA...CCCCTGCCCTCCGATAAACCAACCCCTGGATTTCATTCTGTTC	1531	
319	roGlnGlnIleProAlaValThrLysLysAlaGlnGlyArgThrProAsp	335	
1532	ACTCGGAGATT.....	1542	

OM of: US-09-994-064-11 to: EST:* out_format : pfs
Date: Sep 29, 2002 10:29 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09994064/runat_27092002.185413_13619/app_query.fasta_1.499
-DB=EST -QFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.500 -FGAPOP=6.000
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09994064_ECGN1_1_4874
-NCPU=6 -ICPU=3 -LONGLOG -DEV=US09994064_ECGN1_1_4874
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-994-064-11
Query length: 434
Database: EST:*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 1813.990000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_htc:AK004961	+	117.00	169.00	3.63	3012	AK004961 Mus musculus adult ma
gb_est2:BG467526	+	115.00	182.79	0.6187	661	BG467526 Na.L3_04H04_SAC Necato
gb_est2:BI405354	+	114.00	179.99	0.8865	727	BI405354 117B12 Mature tuber 1a
gb_est2:BG468291	-	114.00	174.95	1.69	1132	BG468291 602509889F1 NIH_MGC_1
gb_est2:BI903542	+	113.50	177.43	1.23	844	BI903542 603168205F1 NCI_CGAP_1
gb_est2:BI948443	+	113.00	181.79	0.7033	533	BI948443 HVSME10009J10f Hordeum
gb_est2:BG599972	+	113.00	177.84	1.17	755	BG599972 EST504867 CSTS Solanum
gb_est2:BE499479	+	111.50	177.94	1.15	596	BE499479 WHE0974_A07_B14ZS Whea
gb_est2:BI684767	+	111.50	174.61	1.77	799	BI684767 603307134F1 NCI_CGAP_2
gb_est2:BG480929	+	111.50	167.46	4.42	1499	BG480929 602530078F1 NIH_MGC_2
gb_est1:AL503461	+	111.00	175.25	1.63	700	AL503461 AL503461 Hordeum vulg
gb_est2:BE412830	+	111.00	171.73	2.56	954	BE412830 MCG008.B08R990625 ITFC
gb_gss:CNS06X9U	-	111.00	171.72	2.56	955	AL419464 T3 end of clone AX0AAO
gb_est2:BE599504	+	110.00	176.63	1.36	533	BE599504 P11_88_C12.bl_A002 Pat
gb_est2:BE599342	+	110.00	176.35	1.41	546	BE599342 P11_88_C12.g1_A002 Pat
gb_est2:BE363291	+	110.00	175.47	1.58	590	BE363291 W51_61_C02.g1_A002 Wat
gb_gss:AQ639480	-	110.00	174.37	1.82	650	AQ639480 927P1-1267.TP 927P1 Tr
gb_gss:AG136530	+	110.00	168.42	3.91	1097	AG136530 Pan troglodytes DNA,
gb_est2:BE797927	-	109.50	168.62	3.81	999	BE797927 601586289F1 NIH_MGC_7
gb_est1:AL506713	+	108.50	170.94	2.83	700	AL506713 AL506713 Hordeum vulg
gb_est2:BF781936	+	107.50	169.25	3.51	698	BF781936 602105889F1 NCI_CGAP_1
gb_est1:AU033360	-	107.00	170.19	3.12	596	AU033360 AU033360 Dictyostelium
gb_gss:CNS044WA	+	106.00	170.40	3.03	503	AL274627 Tetraodon nigroviridis
gb_gss:CNS071BM	+	106.00	162.70	8.14	990	AL424712 T3 end of clone XAZ0AA
gb_est2:BF137567	+	106.00	157.61	15.63	1549	BF137567 601780583F1 NCI_CGAP_1
gb_est1:AL633932	+	105.50	168.73	3.76	540	AL633932 AL633932 XGC-gastrula
gb_est1:AV595249	+	105.50	167.72	4.28	590	AV595249 AV595249 Bos taurus ca
gb_est2:BI888412	+	105.50	163.94	6.95	823	BI888412 ZF637-2-000167 Zebraf
gb_htc:AK011186	+	105.50	160.13	40.83	2774	AK011186 Mus musculus 10 days
gb_est2:BI877213	+	105.00	166.50	5.00	609	BI877213 f182c05.y1 Sugano Kwa
gb_est2:BG331584	+	105.00	165.30	5.83	677	BG331584 602433271F1 NIH_MGC_18
gb_est1:AV400134	+	104.00	169.19	3.54	413	AV400134 AV400134 Bombyx mori b
gb_gss:TA169C02Q	-	104.00	165.51	5.68	571	AL473402 T. brucei sheared genc
gb_est1:AU004829	+	104.00	163.85	7.03	661	AU004829 AU004829 Bombyx mori b
gb_est2:BE255977	-	104.00	163.66	7.20	672	BE255977 601113371F1 NIH_MGC_16
gb_est1:AV404720	+	104.00	163.32	7.51	692	AV404720 AV404720 Bombyx mori b
gb_est1:AU004670	+	104.00	162.27	8.60	759	AU004670 AU004670 Bombyx mori b
gb_est1:AU004186	+	104.00	161.93	8.98	782	AU004186 AU004186 Bombyx mori b
gb_est2:BG478822	+	104.00	161.03	10.09	847	BG478822 602525372F1 NIH_MGC_20
gb_gss:CNS0717H	-	104.00	159.82	11.78	942	AL424563 T3 end of clone XAZ0AA

gb_est2:BG366589	+	103.50	166.16	5.22	500	BG366589 HVSME10007123f Hord
gb_est1:AW950656	+	103.50	163.77	7.10	617	AW950656 EST362726 MAGE rese
gb_est2:BF781801	+	103.50	157.31	16.25	1089	BF781801 602107335F1 NCI_CG
gb_est2:BG289060	+	103.00	159.55	12.19	829	BG289060 602383934F1 NIH_MGC
gb_est2:BI903509	+	103.00	159.49	12.28	833	BI903509 603168013F1 NCI_CGA

seq_name: gb_htc:AK004961

seq_documentation_block:

LOCUS AK004961 3012 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300010N03:transcription factor UBF, full insert sequence.

ACCESSION AK004961 GI:12836541
VERSION AK004961.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library clone:1300010N03.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wachihi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schiml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

2171 CTCTTCTGAGGATGGGGGACTCTTCTGAGTCCAGCAGTGAAGATGAAA 2220

369 luValProGluAspThrGluHisAspAspProAsnSerAspProAspTyr 385
::: ||| |||::: ||| ||| : : : : : ||| |||
2221 GCGAGGATGGGAT...GAGATGATGATGACGACGATGAAGATGAC 2267

386 TyrAsnAsp 388
::: |||
2268 GAAGATGAC 2276

seq_name: gb_est2:BG467526

seq_documentation_block:
LOCUS BG467526 661 bp mRNA linear EST 21-MAR-2001
DEFINITION Na_L3_04H04_SAC Necator americanus (parasitic nematode) L3 Necator americanus CDNA clone Na_L3_04H04 5' similar to pIR1T46707
proteophosphoglycan, membrane-associated - Imported - leishmania,
mRNA sequence.
ACCESSION BG467526
VERSION BG467526.1 GI:13418130
KEYWORDS EST.
SOURCE Necator americanus.
ORGANISM Necator americanus.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
Ancylostomatidae; Ancylostomatidae; Bunostominae; Necator.
REFERENCE 1 (bases 1 to 661)
AUTHORS Blaxter,M.L., Parkinson,J., Whitton,C., Daub,J., Guiliano,D., Hall
N., Quayle,M. and Barrell,B.
TITLE Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared for Dr. David Pritchard University of
Nottingham Sequencing was performed by Claire Whitton ICAPB,
Edinburgh
PCR Primers
FORWARD: SAC
BACKWARD: T7PL
Plate: 04 row: H column: 04
Seq primer: SAC
High quality sequence stop: 486.
Location/Qualifiers
1..661
/organism="Necator americanus"
/db_xref="taxon:51031"
/clone_lib="Na_L3_04H04"
/clone_lib="Necator americanus (parasitic nematode) L3"
/sex="Mixed"
/dev_stage="L3"
/note="Vector: PCMV-PCR vector; Site_1: EcoRI (5'end);
Site_2: XhoI (3'end); Necator americanus is a human
hookworm, responsible for debilitating anaemia. The
library was constructed by David Pritchard (University of
Nottingham, UK) from mRNA from N.americanus L3 stage
(Nottingham strain) maintained in hamsters."

BASE COUNT 203 a 203 c 142 g 113 t
ORIGIN

alignment_scores:
Quality: 115.00 Length: 115
Ratio: 1.885 Gaps: 5
Percent Similarity: 53.043 Percent Identity: 32.174

alignment_block:
US-09-994-064-11 x BG467526 ..

Align seg 1/1 to: BG467526 from: 1 to: 661

309 SerAlaPro.....AspProArgProAspSerValProGluIlePr 323
||||| |||::: ||| ||| ||| ||| ||| ||| ||| |||
313 TCTGCTCATCGCAGGACGACGACCTGATCAAGTTCTCAGAGGCCCC 362

323 oAlaValThrLysLysAlaGluGlyArgThrProAspAlaGluSerSerg 340
||||| :::: ||| : : : : : ||| :
363 TGCACAAATGAACCTGCTCTGAGGAATCCGCACAAAGTGAACCTTAC 412

340 luLysLysAlaProProGluAspSer..GluAspAspMetGlnAlaGlu 355
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
413 AGGAAGCAGCTCCCTCAAAATGCCGTATCAGAGATGCTCCAGCTGAA 462

356 AlaSerGlyGluAsnPro..... 361
||| ||| |||
463 GCTCCTCAAGAGCGCCAGTACAGGCTGAACCAGACATGCTTCCGAATC 512

362AlaAlaLeuProGluAspAspGluValProGluAspThrG 375
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
513 CTTGACCACTGCCGAGGCACCTACCGAC.....GCAGAACTACAC 553

375 luHisAspAspProAsnSerAspProAspTyrTyrAsnAspMetProAla 391
||| |||::: ||| ||| : : : : : ||| |||
554 CACAC.....CCAGATCTGCTCTCCAGCTCAAGAACTTCTTGCTGCC 597

392 ValIleProValGluGluThrThrLysSerSerAsnAlaValSer 406
::: ||| : : : : : ||| ||| |||
598 GAAGTTACACAGAGAGCCCGACGCTGAGTCGGAACCTGCACCATCA 642

seq_name: gb_est2:BI405354

seq_documentation_block:
LOCUS BI405354 727 bp mRNA linear EST 14-AUG-2001
DEFINITION 117B12 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
sequence.
ACCESSION BI405354
VERSION BI405354.1 GI:15184768
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 727)
AUTHORS Nielsen,K.L., Crookshanks,M., Emmersen,J. and Welinder,K.G.
TITLE EST-sequencing of mature potato tuber (Var. Kurasa)
JOURNAL Unpublished (2000)
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgy@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 727
POLY-A=NO.
Location/Qualifiers
1..727
/organism="Solanum tuberosum"
/cultivar="Field grown Kurasa"
/db_xref="taxon:4113"
/clone_lib="Mature tuber lambda ZAP"
/tissue_type="Tuber"
/note="Vector: lambda ZAP"

BASE COUNT 238 a 127 c 198 g 164 t
ORIGIN

alignment_scores:
Quality: 114.00 Length: 268
Ratio: 0.927 Gaps: 10


```

Percent Similarity: 45.896      Percent Identity: 22.761

alignment_block:
US-09-994-064-11 x B1405354      ..

Align seg 1/1  to: B1405354  from: 1  to: 727

134 SerAlaAlaLysGluLysGluLysGlyGlyProPheGluAlaSerValVa 150
   |||||:::|||||:::|||||  |||  ::|||  |||
   TCTGCTTCATCCGAGCCAGCATCTGTATTCCGGGATCTGCCGTAGTC.. 138

150 lTrpPheTyrValIleLysGlyAspSpGlyLysPlystYrCysProI 167
   139 .....TGTGTAA 145

167 lEtyrArgLysGluTyrArgGluCysGlyAspValGlnLeuSerGlu 183
   ||  |||||  |||||  |||||  |||  |||||:::
   TTGTA....CCGTACCGTACTGATGAGACAAGCAGAGGCTCAGTCAG 189

184 CysAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerT 200
   |||  |||:::  |||  |||
   TGTGCCATGATTAGACAGAGTATG..... 214

200 hrLeuValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAla 216
   215 .....CTA 217

217 AlaLeuSerGlyGlnTyrIleLeuThrLeuLysIleGlyArgPheAlaGl 233
   ::  ::|||  |||:::||||  |||:::  :::::
   ACTTGC GCCGTCATCATCTGATAGCTTATTATTCTTCGCTGGTTAA 267

218 ACTTGC GCCGTCATCATCTGATAGCTTATTATTCTTCGCTGGTTAA 267

223 nThrAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlyS 250
   :|||  :::::|||||  |
   AACTCAGATCCTCACTCTC.....ACCCAGG 293

250 ergLLeuAsnPheLeuProSerLysCysTrpThrThrGluGlnTyrGln 266
   |||:::||||  ::|||:::  |||:::||||
   TGTACTTGAAGCTACTCGAACAAAGGCTAAAGGCTGAAGAACATCAT 343

294 TGTACTTGAAGCTACTCGAACAAAGGCTAAAGGCTGAAGAACATCAT 343

267 ThrGlyPheGlnGlyGlnHisLeuTyrProIleAlaAspThrAsn..Th 282
   :::  :::::  |||  |||  |||  |||
   ATAATCTCCCAAGCAACACTCTTGATGCCGAAGAAGACGGCAATGAGAC 393

344 ATAATCTCCCAAGCAACACTCTTGATGCCGAAGAAGACGGCAATGAGAC 393

282 rArgHisAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgT 299
   |  |||||  ::|||  :::
   GGGGATGATGATGATGATGATTAAGGTGAAGACTTGTCATCCGAAGATG 443

394 GGGGATGATGATGATGATGATTAAGGTGAAGACTTGTCATCCGAAGATG 443

299 rPAsnAsnLeuLeuArgLysLysAsnProSerAlaProAspProArgPro 315
   ::|||:::  :::::|||||:::  |||  :::
   GAGGAATGTG....AACATAACCCAAACCAATGATAGCAACTCC 487

444 GAGGAATGTG....AACATAACCCAAACCAATGATAGCAACTCC 487

316 AspSerValProGlnGluIleProAlaValThrLysLysAlaGluGlyAr 332
   :::  |||:::||||  ::|||:::||||
   AAGAAACATCCTAGAGAA.....GAGCAGACGGTGC 519

488 AAGAAACATCCTAGAGAA.....GAGCAGACGGTGC 519

332 g..... 332

520 TGAAGAGAAATGCTGAAGATGATGATGATGATGAGGGGTGTTGAAG 569

333 ..ThrProAspAlaGluSerSerGluLysLysAlaProProGluAspSer 348
   ::|||  |||:::||||  |||:::||||
   ACAGTCCAGATGATGAAGATTAATGATAA.....CCTGATGAAGAA 610

570 ACAGTCCAGATGATGAAGATTAATGATAA.....CCTGATGAAGAA 610

349 GluAspAspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuPr 365
   ::|||  :::::  ::|||  :::
   GATGATGACGAAGATGGAACCTGATGAAGGCGAGGAGATACTGTGGAAGA 660

611 GATGATGACGAAGATGGAACCTGATGAAGGCGAGGAGATACTGTGGAAGA 660

365 oGluAspAspGluValProGluAspThrGluHisAspAspProAsnSerA 382
   |||||  |||  |||  ::|||  :::::
   AGAGGACGATGAAGAGCAAGAGCAAGAGCAAGATGAAGTGAAGAAAG 710

661 AGAGGACGATGAAGAGCAAGAGCAAGAGCAAGATGAAGTGAAGAAAG 710

382 sp 382
::

```

```

711 AA 712
seq_name: gb_est2:BG468291
seq_documentation_block:
LOCUS BG468291 1132 bp mRNA linear EST 21-MAR-2001
DEFINITION 602509889F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4643999 5',
ACCESSION BG468291
VERSION BG468291.1 GI:13400561
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1132)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM1415 row: a column: 24
High quality sequence stop: 528.
FEATURES
Source
Location/Qualifiers
1..1132
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4643999"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: POTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 400 a 314 c 268 g 150 t
ORIGIN
alignment_scores:
Quality: 114.00 Length: 356
Ratio: 0.750 Gaps: 17
Percent Similarity: 42.697 Percent Identity: 23.034
alignment_block:
US-09-994-064-11 x BG468291/rev ..
Align seg 1/1 to reverse of: BG468291 from: 1 to: 1132
25 CysGlyGlyLysArgCysCysSerGlyAlaAlaValPheThrLeuPheTr 41
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
812 TGCTTCCCTTTTCGTTTCGTTCTGTTGGGCTGGTTGTCCTGTACCTGTTGTG 763
41 P..... 41
|
762 GGGTTTGTGTTGTTGCTAGTCTGTTGGATTGATTTTCGATTGAT 713
42 .....ThrcysValArgIleMetArgGluHisIleCysPheVal 54
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
712 TGCTTGCATGTTCTGCTGCTAATCTTCGAGTCTGCTGCTGCTTTTGT 663
55 ArgAsnAlaMetAspArgHisLeuPheLeuArgAsnAlaPheTrThrII 71

```



```

662 GGGCGCTCGA.....TTTTGGCGCATCCTCTTTGGCGAGT 628
71 eval...leuLeuSerSerPheAlaSerGlnSerThrAlaIaValThr 87
: |||||:::|::| |||:::|::|
627 AGGTAACCTGTTGTCTCAGGATTTCTG.....ACGTTCTGTGTG... 591
87 yrasPTYrIleLeuGlyArgArgAlaLeuAspAlaLeuThrIleProAla 103
591 ..... 591
104 valGlyProTYrAsnArgTYrLeuThrArgValSerArgGlyCysAspVa 120
590 .....ATATCCAGGTGATCCGAGGATTTTCT 562
120 lValGluLeuAsnProIleSerAsnValAlaAspMetIleSerAlaAla 137
: ::|||::: |||:::|::|
561 GCGTCGCTTACACGCTCAGTTTGCCATGTGATCTCTGCTGAGCCTCACCAT 512
137 ysgIuLysGluLysGlyGlyProPheGluAlaSerValValTyrPheTYr 153
511 GTATGGAGCTAAATTGTTGGTTTGTGATTGTGACCTC.....TTTCAG 468
154 valIleLysGlyAspAspGlyIuAspLysTYrCysPro.IleTYrArgL 170
|||:::| |||:::| |||:::|
467 TCGATAGGTTCTCCGGATGACATTCAGTGATTTGCCCATTAAGTCTCAG. 419
170 ysgIuTYrArgGluCysGlyAspValGlnLeuLeuSerGluCysAlaVal 186
418 .....GACCTACAGGAACCTTTGATTGGTGACCA 389
187 GlnSerAlaGlnMetTyrAlaVal.....AspTYr.ValProSerThrL 201
:::|:::|:::|:::|:::|:::|
388 GAATCCAGTACTTCTGCGCTCTAACCAGCCGTTCTGTACCATTCATAG 339
201 euValSerArgAsnGlyAlaGlyLeuThrIlePheSerProThrAlaAla 217
:::|:::| |||||:::| |||||
338 CAATAGTC.....CTTACTGTGCTCACC..... 313
218 leuSerGlyGlnTYrLeuLeuThrleuLysIleGlyArgPheAlaGlnTh 234
312 .....AAATGCTGGGCCACTTCCAAAAC 290
234 rAlaLeuValThrLeuGluValAsnAspArgCysLeuLysIleGlySerG 251
|:::| |||||:::|:::|:::|:::|:::|
289 CAGTCTGGTCTCCCTGCCCTTGCACTTCCAGGGCATTTAGAAATGGTGTA 240
251 InleuAsnPheLeuProSerLysCysTrpThrThrGluGlnTYrGlnThr 267
:::|:::| |||||:::| |||||:::|
239 GT.....CCCTCATCAAACTGAGCTCACC..... 214
268 GlyPheGlnGlyGluHisLeuTYrProIleAlaAspThrAsnThrArgHi 284
213 .....ACTGCCCAATGACCGCCACG..... 193
284 salaAspAspValTYrArgGlyTYrGluAspIleLeuGln.ArgTrpAsn 300
192 .....ATGCGCCCGGTGGCGG 177
301 AsnLeuLeuArgLysLysAsnProSerAlaPro..AspProArgProAs 316
:::| |||||:::|:::| |||||:::| ||
176 CGCCTGCTTTTGGCGAAGGAGATGTTTGGCGCGCATAGTCCCTGACAGGA 127
316 pSerValProGlnGluIleProAlaValThrLysLysAlaGluGlyArgT 333
|:::| |||||:::| |||||:::|:::|
126 TGGACCGCGCGTGGAGCGCGCCGACAGTAAAGAGCTGAGCTGGGGGACGGA 77
333 hrProAspAlaGluSerSerGluLysLysAlaProProGluAspSerGlu 349
76 CGCTGAAGGGGTGAGTCTCCGCAAGGCCCCGAGGCGCGAGCAGCGGCCA 27
350 AspAspMetGln 353
||| |||

```

26 CCCGACCCACAA 15

seq_name: gb_est2:BI903542

seq_documentation_block:

LOCUS BI903542 844 bp mRNA linear EST 16-OCT-2001

DEFINITION 603168205F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5256289 5',

mRNA sequence.

ACCESSION BI903542

VERSION BI903542.1 GI:16165489

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 844)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM11647 row: j column: 02
High quality sequence start: 83
High quality sequence stop: 836.

FEATURES

Source

```

BASE COUNT      259 a      205 c      223 g      155 t      2 others
ORIGIN
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5256289"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGACGGCCGCTGTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaudo. "
```

alignment_scores:

Quality:	113.50	Length:	190
Ratio:	1.124	Gaps:	8
Percent Similarity:	53.158	Percent Identity:	26.316

alignment_block:

Align seg 1/1 to: BI903542 from: 1 to: 844

```

270 gln gly u hi s leu tyr pro ..... i le a la asp thr as 281
    ||||| : : : : ||| : : : :
222 caa ggt aac tca tcc atc cta aag cca tcc aca tta tag ta gaa gca 271
    ||||| : : : : ||| : : : :
281 n th r a g h i s a l a s p a s p v a l t y r a r g l y t y r g l u a s p i l e ..... 295
    : ||||| : : : : |||||
272 g c t g a t c c a c g c g g a c a g t t a g g t g a a a g c a g a g a t a c c a c a t g c 321

```

296 ..LeuGlnArgTrpAsnLeuLeuArgLysLysAsnProSerAla... 310
:::||||| ||| ||::: ||:::|
322 CTATACGTCGAGCCCGTAATTCTACCCGGGAAACTCCACCCAAAAGCAAA 371
311ProAspProArgProAspSerValProGlnG1 321
||:::|||||::: ||:::|
372 CTTGCTGAAGGGGAGAGAAAGAAAACCAAGACAGATGAGATTGAGAGGA 421
321 uileProAlaValThrLysLysAlaGluGlyArgThrProAspAlaGluS 338
| :::::|:::|:::|:::|:::|:::|:::|:::|
422 ATCTATCTCTACTAGTAAGAGAACAGAGATGAGACTCCACCTGCTACAT 471
338 erSerGluLysLysAlaPro.....ProGluAspSerGluAspAsp 351
|||||||::: ||| |||||:::|:::|:::|:::|
472 CCAGTGAGGCGAGAGCAAGCCCAAGGGGAGCGCTGAGAGTGAGAGAGAA 521
352 MetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAspAs 368
::: ||: ||:::|:::|:::|:::|:::|:::|:::|
522 GAGAACCAACAGCAAGTCTGCTGAGGAACCCAAAGATGAGAAAGATCA 571
368 p.....GluValProGluAspThrGluHisAspAsp..... 378
: ||| ||:::|:::|:::|:::|:::|:::|
572 GTCTAAAGAAAAGAGAGAAAGTGAAGAAAGCATTCCTGCTGGCGG 621
379ProAsnSerAspProAspPtyrTyrAsnAspMetPro 390
||| ||:::|:::|:::|:::|:::|:::|
622 ACTCTCTCTGCGCAGCAGCTAGCCAGGCCAGACAAACCCCATG... 668
391 AlavalIleProValGluGluThrThrLysSerSerAsnAlaValSerMe 407
::: ||: ||:::|:::|:::|:::|:::|:::|
669 GGCTTCCTCCCC.....ACGGCCAAGATGAGCGCAATCTTAAGTGA 709
407 tProIlePheAlaAlaPheValAlaCysAlaValAlaLeuValGlyLeuL 424
||| ||:::|:::|:::|:::|:::|:::|:::|
710 GGCCATTAAAGGCATGCTTCCAGAAAGACTGGCGCGCTGGTGGCATTC 759
424 euValTrpSerIleValLys 430
::: ||| |||
760 GAAAGTACATCATTCATTAAG 779

seq_name: gb_est2:BI948443

seq_documentation_block:

LOCUS BI948443 533 bp mRNA linear EST 19-OCT-2001
DEFINITION HVSME10009J10f Hordeum vulgare spike EST library HVCDNA0012
(Fusarium infected) Hordeum vulgare cDNA clone HVSME10009J10f, mRNA
sequence.

ACCESSION BI948443
VERSION BI948443
KEYWORDS BI948443.1 GI:16288851
EST.

SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

REFERENCE 1 (bases 1 to 533)
AUTHORS Wing,R., Muehlbauer,G.J., Close,T.J., Kleinhofs,A., Wise,R., Heinen
,S., Begum,D., Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T.,
Simmons,J., Fenton,R.D., Malatrasi,M., Choi,D.W., Oates,R. and Main

TITLE 'D.
COMMENT Development of a genetically and physically anchored EST resource
for barley genomics: Fusarium infected Morex spike cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 123
Seq primer: AATTAAACCTCACTAAAGG
High quality sequence start: 2

FEATURES High quality sequence stop: 520.
location/Qualifiers
source 1..533
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSME10009J10f"
/clone_lib="Hordeum vulgare spike EST library HVCDNA0012
(Fusarium infected)"
/tissue_type="Spike"
/lab_host="TJC121"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Plants were grown at the University of Minnesota in
the GJ Muehlbauer lab; spikes were harvested and snap
frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium
graminearum inoculation (Heinen). In the TJ Close lab at
the University of California, Riverside, total RNA was
prepared from each sample pool, equal quantities of all
eight RNA pools were combined, poly(A) RNA was purified
from the mixture, one primary unamplified cDNA library was
made, and 1 million pfu were in vivo excised to give
pBluescript SK(-) cDNA phagemids (Choi, Fenton, Malatrasi
, Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

BASE COUNT 183 a 141 c 160 g 49 t
ORIGIN

alignment_scores:
Quality: 113.00 Length: 111
Ratio: 1.852 Gaps: 4
Percent Similarity: 54.955 Percent Identity: 28.829

alignment_block:
US-09-994-064-11 x BI948443 ..

Align seg 1/1 to: BI948443 from: 1 to: 533

305 LysLysAsnProSerAlaProAspProArgProAspSerVal..ProG1 320
:::|||||:::|::: ||| |||||:::|:::|:::|
29 GAAGAGAAAGCGCTGACTGACGACAGAGAGGCCGATGAGTTAAAGAAAGA 78
320 nGluIleProAlaValThrLysLysAlaGluGlyArgThrProAspAla. 336
: ||| ||:::|:::|:::|:::|:::|:::|:::|
79 AGAAAGCATGCGCGCGAAAGAAAGCCAGAGCAGCAAAAGCGATGCCA 128
337GluSerSerGluLysLysAlaPro 344
129 GCAACAATGCCGAGAGGCCCAAGCCCGAAGAAAGAAAGAAAGAGCCA 178
345 Pro.....GluAspSerGluAs 350
||| |||||:::|:::|
179 CCTGTCAGACAGAGAGAGAGAGCGAGTCCGAGTCCGAAGATGATGACGA 228
350 pAspMetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluA 367
|||:::|:::|:::|:::|:::|:::|:::|
229 CGACTTTGATTCGGAAGACGACTCA.....GAGG 257
367 sPaspGluValProGluAspThrGluHisAspAspProAsnSerAspPro 383
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

258 AAGACTCTGAGGAAGAGACTCCGAGATGACGAGTCAGAGTCGAGTCC 307
384 AsPTYrTyraSnaspmetProAlaValIlePro 394
   ::::: |||||::: |||
308 GAGCAGGAGGAGGAGAGACCTGCCACTCCACC 340

seq_name: gb_est2:BG599972

seq_documentation_block:
LOCUS      BG599972                755 bp      mRNA      linear      EST 12-APR-2001
DEFINITION EST504867 cSTS Solanum tuberosum cDNA clone cSTS26J20 5' sequence,
            mRNA sequence.
ACCESSION  BG599972
VERSION    BG599972.1  GI:13617108
KEYWORDS   EST.
SOURCE      potato.
            Solanum tuberosum
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE  1 (bases 1 to 755)
AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE     Generations of ESTs from sprouting potato eyes
JOURNAL   Unpublished (2000)
COMMENT   Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com
            Seq primer: M13F-R.

FEATURES
            Location/Qualifiers
                source
                    1..755
                        /organism="Solanum tuberosum"
                        /cultivar="Kennebec"
                        /db_xref="taxon:4113"
                        /clone="cSTS26J20"
                        /clone_lib="cSTS"
                        /tissue_type="sprouting eyes from tubers"
                        /dev_stage="12-14 weeks post harvest"
                        /lab_host="SOLR"
                        /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                        XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                        taken from tubers. The tubers were incubated at 26C in the
                        dark for 2-3 weeks prior to sprouting. The eyes were
                        frozen in liquid nitrogen immediately upon removal from
                        tubers."
BASE COUNT  254 a      134 c      195 g      172 t
ORIGIN
alignment_scores:
    Quality: 113.00      Length: 264
    Ratio: 0.942      Gaps: 10
Percent Similarity: 45.455      Percent Identity: 23.106

alignment_block:
US-09-994-064-11 x BG599972 ..

Align seg 1/1 to: BG599972 from: 1 to: 755

134 serAlaAlaIySgIuTyraGluCySgIyProPhelGluAlaSerValVa 150
|||||:::|||||::: ||| ::||| |||
51 TCTGCTTCATCGAGCGACGATCTGTATTTCGCGGATCTGCGCTAGTC.. 98

150 lTrpHetYrValIleIySgIyAspAspGlyGluAspIySTyrcYsProI 167
|||||
99 .....TGTGTAA 105

167 lETyrrArGlySgluTyraRgGluCySgIyAspValGlnLeuLeuSerGlu 183
||||| ||||| ||||| |||||
106 TTGTAA.....CCGTACCGTACTGATGGAGACAAGCAGGCGCTCAGTCAG 149

184 Cys.AlavaGlnSerAlaGlnMetTrpAlaValaIaSPtyrValProSerT 200

```

```

150 TGTGCCATGATTAGACAGAGTATG..... 174
175 .....,CTA 177
200 hrLeuValSerArgsnGLyAlaGLyLeuThrIlePheSerProThrAla 216
217 AlaLeuSerGLyGLnTyrrLeuLeuthrLeuLysIleGLyargPheAlaGL 233
::: ||||| | | : : : : : ||| : : : : :
178 ACTTGCGCCGGTCATCATCAGTAGCTTTATTATTCTTCGCTGTAA 227
233 nThrAlaLeuValThrLeuGLuValAsnAsparGCysLeuLysIleGLys 250
: || : : : : |||
228 AACTCAGATCCTCACTCTC..... ACCCAGG 253
250 erGLnLeuAsnPheLeuProserLysCystrrThrThrGLnGLnTyrrGLn 266
||| : : : : ||| : : : : : ||| : : : : :
254 TGTAAGCTTGAAGCTACTCGAACAAAGGCTAAAAGGCTGAAGAATCAT 303
267 ThrGLyPheGLnGLyGLnHisLeuTyrrProIleAlaAspThrAsn...Th 282
::: : : : : ||| ||| |||
304 ATATCTTCCCCAAGCAAACCTCTTGATGCCGAAGAAGACGGCATGAGAC 353
282 rArghisAlaAspAspValTyrrArgGLyTYrrGLuAspIleLeuGLnArgT 299
| ||||| : : ||| : : :
354 GGGGATGATGATGATGATGATAAGGTGAAGACTTGTATCCGAGATG 403
299 rPasAsnLeuLeuArgLysLysAsnProSerAlaProAspProArgPro 315
::: ||| : : : : : ||| : : :
404 GAGCAATGTG.... AACAAATACCACAACCAATGATAGCACTCC 447
316 AspSerValProGLnGLuIleProAlaValThrLysLysAlaGLnGLyAr 332
::: ||| : : ||| : : : : |||
448 AAGAAACATCCTTAGAGAA..... GAGCAGACGGTGC 479
332 g..... 332
480 TGAAGCAATGCTGAAGATGATGATGATGATGAGCGGTGCTGAAG 529
333 .ThrProAspAlaGLuSerSerGLuLysLysAlaProProGLuAspSer 348
::: ||||| | | : : : : : ||| : : : : :
530 ACAGTCCAGATGATGAAGATAATGATAAA..... CCTGATGAAGAA 570
349 GluAspAspMetGLnAlaGLuAlaSerGLyGLuAsnProAlaAlaLeuPr 365
::: ||||| : : : : : ||| : : : : :
571 GATGATGACGAAGATGACCTGATGAAGCGCAGGAAGATACTGTGAAGA 620
365 OGluAspAspGLuValProGLuAspThrGLnHisAspAsp 378
||||||| | | : : |||
621 AGAGCAGCATGAAGAGGAAGAGGAAGATGAAGAT 660

seq_name: gb_est2:BE499479

seq_documentation_block:
LOCUS BE499479 596 bp mRNA linear EST 04-AUG-2000
DEFINITION WHE0974_A07_B14ZS wheat pre-anthesis spike cDNA library Trilicium
aestivum CDNA clone WHE0974_A07_B14, mRNA sequence.
ACCESSION BE499479
VERSION BE499479.1 GI:9698096
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
; Triticeae; Triticum.
1 (bases 1 to 596)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
COMMENT Contact: Olin Anderson
```

US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

Email: oandersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: StrataGene SK primer.

FEATURES

source

location/Qualifiers
1.596
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone_lib="WHE0974_A07_B14"
/tissue_type="Wheat pre-anthesis spike cDNA library"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give Bluescript
phagemids in the T7 Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

BASE COUNT 164 a 155 c 187 g 90 t
ORIGIN

alignment_scores:

Quality: 111.50 Length: 136
Ratio: 1.689 Gaps: 6
Percent Similarity: 48.529 Percent Identity: 30.882

alignment_block:

US-09-994-064-11 x BE499479 ..

Align seg 1/1 to: BE499479 from: 1 to: 596

284 HisAlaAspAspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAs 300
||| ::|||::|
60 CACAAGGAAGATGTCAAACACGCTGAGTGTGTCCTCCGCA..... 104
300 nasLeuLeuArgLysLysAsnProSerAlaProAspProArgProAsps 317
105CCAGCGCGGAGAGACAGAAGAAAGATG 132
317 erValProGlnGluLe.....ProAlaValThrLysLysAlaGlu 330
||| ||| ||| ||| ||| |||
133 TCAAACCACTGAGACTGATGCCCTCCACACCAACA.....GTGGA 176
331 GlyArgThrProAspAlaGluSerSerGluLysLysAlaProPro..... 345
::| ::| ::| ::| ::| ::|
177 GACAAGAAGAGATGTCAAGCCAGCTGAGCCGATGCCCTCCGCAACC 226
346GluAspSerGluAspAspMetGlnAlaGluAlaSerGlyGlu 360
||||| ::| ||| ||| ||| |||
227 AGCGGTGAAGACAAGAGACGCTTAAGCCAGCTGAGGCTGATGCCCT. 275
360 snProAlaAlaLeuProGluAspAspGluValProGluAspThrGluHis 376
||||| ::| ||| ||| ||| |||
276 ..CCGGCACCAAGCGGTTGAAGACAGAAGACGCTAAGCCAGCTGAGGCT 323
377 AspAspProAsnSerAspProAspTyrTyrAsnAspMetProAlaValIle 393
||| ::| ::| ::| ::| ::|
324 GATGCTGCTCCGCA.....CCAGCGGTGGA 349
393 eProValGluGluThrThrLysSerSerAsnAlaValSerMetProIleP 410
::| ||| ::| ||| ||| ||| |||

350 AGACAAGAGAAGATGTCATCATCTGAGGCTGATGCCGCTCTGTGG 399
410 heAlaAla 412
::| ::|
400 TGAGTTTCG 407

seq_name: gb_est2:BI684767

seq_documentation_block:

LOCUS BI684767 799 bp mRNA linear EST 18-SEP-2001
DEFINITION 603307134F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5342936 5',
mRNA sequence.
ACCESSION BI684767 GI:15647395
VERSION BI684767.1 GI:15647395
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 799)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11871 row: d column: 09
High quality sequence stop: 797.

FEATURES

source

location/Qualifiers
1..799
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5342936"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 246 a 201 c 213 g 139 t
ORIGIN

alignment_scores:

Quality: 111.50 Length: 195
Ratio: 1.115 Gaps: 9
Percent Similarity: 51.282 Percent Identity: 25.128

alignment_block:

US-09-994-064-11 x BI684767 ..

Align seg 1/1 to: BI684767 from: 1 to: 799

270 GlnGlyGluHisLeuTyrPro.....IleAlaAspThrAs 281
||||| ::| ||| ||| ||| |||
185 CAAGGTGACTCATCTCCTAAGGCACTCCCACTTATAGTAGGAGCGCA 234
281 nThrArgHisAlaAspAspValTyrArgGlyTyrGluAspIle..... 295
::| ::| ::| ::| ::|
235 GGTGATCCACGCGGACAGTAGGTGAGAAAGACAGAGATACCAACATGC 284
296 ..LeuGlnArgTrpAsnAsnLeuLeuArgLysLysAsnProSerAla... 310
::| ||| ::| ||| ||| ||| |||


```
285 CTATACGTCGAGCCGTGAATTCTACCCGGGAAACTCCACCCAAAGCAA 334
311 .....ProAspProArgProAspServalProGlnG1 321
335 CTTCGTGAAGGGGAGGAAGAAAACCAGAACCATGGAAGTTCAGAGGA 384
321 uileProAlaValThrLysLysAlaGluGlyArgThrProAspAlaGluS 338
385 ATCTAATCTCTACTGTAGAGAGAGAGAGAAATGAGACTCCACCTGTACAT 434
338 exSerGluLysLysAlaPro.....ProGluAspSerGluAspAsp 351
435 CCAGTCGAGCAGACAGACCCCAAGGGGAGACCTGAGAGTGGAGAGAGAA 484
352 MetGlnAlaGluAlaSerGlyGluAsnProAlaAlaLeuProGluAspAs 368
485 GAGAACACACAGCAAGTCTGCTGAGGAACCCAAAGAGATGAGAGATCA 534
368 p.....GluValProGluAspThrGluHisAspAsp..... 378
535 GTCTTAAGAACAGAGAGAGAAAGTGAAGAACGATTCCTGCTGGGCG 584
379 .....ProAsnSerAspProAspTyr 385
585 ACTCTCTCTGACCAGCCAGCTAGCCAGGGCCAGAGACATTAACCCATGGC 634
386 Tyr.....AsnAspMetProAlaValIleProValGluGluThrTh 399
635 TTCCTTCCCCACGCGCCAAAGATGAGCAATCTTA.....AC 669
399 rLysSerSerAsnAlaValSerMetProIlePheAlaAlaPheValAlaC 416
670 TGAGGCCATTTAAGGCATGCTTCCAGAGAAGACTGGCGCCTCGGTGGCAA 719
416 ys...AlaValAlaLeuValGlyLeuLeuValTrrp 426
720 TCGAAGATCATCATTAAGTACCCTTCTCTGG 754
```

seq_name: gb_est2:BG480929

seq_documentation_block:

LOCUS BG480929 1499 bp mRNA linear EST 21-MAR-2001
DEFINITION 602530078F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4653474 5',
mRNA sequence.

ACCESSION BG480929

VERSION BG480929.1 GI:13413208

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1499)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1439 row: 1 column: 19

High quality sequence stop: 220.

Location/Qualifiers

FEATURES

Source

1.1499
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4653474"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"

/lab_host="DH10B (phage-resistant)"
/note="Organ: Placenta; Vector: pOTB7; Site_1: xhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected by
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

alignment_scores:

Quality:	111.50	Length:	369
Ratio:	0.697	Gaps:	22
Percent Similarity:	43.360	Percent Identity:	23.577

alignment_block:

US-09-994-064-11 x BG480929 ..

Align seg 1/1 to: BG480929 from: 1 to: 1499

```
25 CysGlyGlyLysArgCysCysSerGlyAlaAlaValPheThrLeuPheTr 41
111 111 111111 111111111111111111111111111111111111
104 TGCCTTGGGACCAGGTGCTTCCGGGGCGAAGCCTTTTGGGACCCTTCTG 153
41 pThrCysValArgIleMetArgGluHisIleCysPheValArgAsnAlaM 58
111 111 1111111111111111111111111111111111111111111
154 GAGC.....TTTGGGAGACGAGCGC 173
58 etAspArgHisLeuPheLeu..ArgAsnAlaPheTrpThrIleValLeu 73
174 TGGAT.....TTCCTGGGGCGGGGATTTTCTGG..... 202
74 LeuSerSerPheAlaSerGlnSerThrAlaAlaValThrTyrAspTyrI1 90
202 ..... 202
90 eLeuGlyArgArgAlaLeuAspAlaLeuThrIleProAlaValGlyProt 107
111111111111111111111111111111111111111111111111111
203 ...GGCAGAAACCTTCTGGGGTCTGGAGC.....CCTT 233
107 yTrAsnArgTyrLeuThrArgValSerArgGlyCysAspValVal..... 121
111 111 1111111111111111111111111111111111111111111
234 TAAAC.....TCCGGGGGGGTAGATCTTATTATGCTC 265
122 ..GluLeuAsnProIleSerAsnValAspAspMetIleSerAlaAlaLy 137
111111111111111111111111111111111111111111111111111
266 TGAGAACTTATAGCAGCATGACAGCAGCAGCCAGCAGCAGCAGCAAG 315
137 sGluLys.....GluLysGlyGlyProPheGluAlaSerValValTrrp 152
111111111111111111111111111111111111111111111111111
316 AGAGAAAGCAAGCCGAGAGAGAAATACCTTTAGATCCCGAGAGTGTGCTTT 365
152 heTyValIleLysGlyAspAspGlyGlu.AsplysTyrCysProIleTy 168
111 111 1111111111111111111111111111111111111111111
366 GTTGGGGAAGAAAGCCTTACTCAGAGAAAGCGTGCCTTTGTGAAGCTT 415
168 rArgLys..GluTyrArgGlyCysGlyAspValGlnLeuLeuSerGluC 184
111111111111111111111111111111111111111111111111111
416 TCATTAACCTCTATACCGTAGA.....ATTAATACCT 447
184 ysaAlaValGlnSerAlaGlnMetTrpAlaValAspTyrValProSerThr 200
111111111111111111111111111111111111111111111111111
448 GTACCTCAATTAATTTC.....TATGGCCCTTACGAA 479
201 LeuValSerArgAsnGly.....AlaGlyLeuThrIlePheSe 213
111 111 1111111111111111111111111111111111111111111
480 TAAATTTCACAAAGAGCAGACTCAAAACATTCTGGGTGTACACTCATGGG 529
213 rProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrLeuLysIleGlyA 230
111 111 1111111111111111111111111111111111111111111
```

```
530 GACTATCACACCTTATCATATCAGCGGACATATACAACCTCCTGTGCTGG 579
230 rgpheaIaGlnThrAlaLeuValThrLeuGluValAsnAspArgCysLeu 246
      ::::::::::: |||||::: |||
580 GCGCCCCAAATTCGTGTGTGTGACTGTGCCCCAAATGT...TGTAGC 625
247 LysIleGlySerGlnLeuAsnPhelLeuProSerLysCys..... 259
      ||| ||||| ::||| ||||| |||
626 AAATGAGAGAGT...GTTAACCGTGGACCATTCACCTTGACACCAAGCGAC 672
260 .....TrrThrThrGlnGlnTyrGln.....Thrg 268
      ||||| ::|||
673 ATAGCTCGGAACAACAATTAACTATCAGTGGGACAGAGACCACTCGAG 722
268 LyrpheGlnGlyGluHisLeuTyrProIleAlaAspThrAsnThr.ArgHi 284
      || ::||| ||| ||||| |||
723 GCGCCGGAGACAACAACCCCTGTGCCGAGAGACGTTACAAACACCTCGC.. 770
284 sAlaAspAspValTyrArgGlyTyr.....GluAspIleLeuGlnArgT 299
      |||||::: ::|::: |||
771 .....AGGGACACCGCCGCTGACCCAGTGGCACACCGT 804
299 rPaAsnAsnLeuLeuArgLysLysAsnProSerAlaProAspProArgPro 315
      ||::||| ||::|::: ::|::: |||
805 GGGGAACAACAAGGACCAAGAGGACTCCGGGTATCACCACTT... 851
316 AspSerValProGlnGluIleProAlaValThrLysLysAlaGluGlyAr 332
      852 .....CACACACTGATGAAGACTGAACAACG 877
332 gThrProAspAlaGluSerSerGlnLysLysAlaProGluAspSerg 349
      ||| ::|::: ::|::: ||||| |||:::
878 GGGTCCCCATACTCAGCAACCAAGAAGAGCCCCCGGAGGATGATC 927
349 LuAspAspMetGlnAla.....GluAlaSerGlyGluAsn 360
      ::|::: |||::|:::
928 AGGGAACCAACAAGAACCCACTAACTAAACAAGAGTCGGGGCGGCAA 977
361 Pro 361
      |||
978 CCC 980

seq_name: gb_est1:AL503461
seq_documentation_block:
LOCUS AL503461 700 bp mRNA linear EST 04-JAN-2001
DEFINITION AL503461 Hordeum vulgare Barke roots Hordeum vulgare cDNA clone
HM02D15T 5', mRNA sequence.
ACCESSION AL503461
VERSION AL503461.1 GI:12029676
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
1 (bases 1 to 700)
REFERENCE 1
AUTHORS Michalek, W., Weschke, W., Pleissner, K.-P. and Graner, A.
TITLE EST sequencing and analysis in barley
JOURNAL Unpublished (2000)
COMMENT Contact: Michalek W
Institute for Plant Genetics and Crop Plant Research
Corrensstr.3, D-06466 Gatersleben, Germany
Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
Seq primer: T3 primer for 5'end.
FEATURES
Source
1..700
Location/Qualifiers
1..700
/organism="Hordeum vulgare"
/cultivar="Barke"
/db_xref="taxon:4513"
/clone="HM02D15T"
/clone_lib="Hordeum vulgare Barke roots"
/tissue_type="roots"
```

```
/lab_host="XL0LR"
/notes="Vector: plasmid pBR-CMV; Site_1: EcoRI; Site_2:
XhoI; mRNA was made from roots of spring barley variety
'Barke', a high quality malting variety. Roots were grown
for two days on filter paper at room temperature Cloning
sites: EcoRI (5'-end of cDNA) and XhoI (3'-end of cDNA).
NOTE: Due to a cloning artefact caused by the kit, in most
cases the EcoRI site is NOT present, as well as the EcoRI
adapter. Average insert size is 1 kb Sequence trimming:
Vector sequences and sequence ends were trimmed from the
5'-and 3'-end until a 50 bp window contains less than two
ambiguities. The maximum length was set to 700 bp"
BASE COUNT 212 a 142 c 192 g 152 t 2 others
ORIGIN

alignment_scores:
Quality: 111.00 Length: 283
Ratio: 0.888 Gaps: 13
Percent Similarity: 44.170 Percent Identity: 22.615

alignment_block:
us-09-994-064-11 x AL503461 ..
Align seg 1/1 to: AL503461 from: 1 to: 700

151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
      ||::| ||| ::||| ||||| |||:::
25 TGGCTGTAT.....GAAGATGGCGAGATGAGACTAAGGAGT 62
167 eTyrArgLysGluTyrArgGluCys.GlyAspValGlnLeuLeuSerglu 183
      :||| ::|::|::: ::|::: |||||:::
63 CTATGTTGCAAAACTAGAAAGAACTGAAAGCTTGCGGCTCTATCGAAA 112
184 CysAlaValGlnSeraIaGlnMet.....TrpAl 193
      |||||::: |||::|:::
113 TGGCTTACAAAGAGTGGACAGAAAGAGTCCGGCTTTGAGCAATTGGTG 162
193 aValAspTyrValProSerThrLeuValSerArgAsnGlyAlaGlyLeuT 210
      :
163 T..... 163

210 hrIlePheSerProThrAlaAlaLeuSergLysGlnTyrLeuThrLeu 226
      ::|::: |||::|:::
164 .....ACTGCATCCGCGAGTTTAAAGAGGCTGCATTGCTGTGACC 205
227 LysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGluValAsnAs 243
      ||| ::|||::|:::
206 AAA.....AGTTGATCACATA.GACATATCAGA 233
243 PArgCysLeuLysIleGlySerGlnLeuAsnPhelLeuProSerLysCysT 260
      ::|::: |||::: ::|:::
234 GAAG...CAAAAGTTGTTAATGAGTGTCTG.....GATGACAGAGACCT 274
260 rPThrThrGlnGlnTyrGlnThrGlyPheGlnGlyGluHisLeuTyrPro 276
      ||| |||::: |||
275 GGCTAATGAGAAAAAGCAGCAG..... 297
277 IleAlaAspThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrG 293
      |||::: |||::: |||
298 ...CAAGATGCTCTACCGAAGCATGCTAATCTGTCTCTCT..GTTTC 341
293 uAspIleLeuGlnArgTyrPaAsnAsnLeuLeuArgLysLysAsnProSera 310
      :||| |||::: ::|::: |||
342 TGACATTAAGAAGAGGCTGAAGCAGCTTGACAGGTTCTGCAAAACCATCA 391
310 laProAspProArgProAspSerValProGlnGluIlePro...AlaVal 325
      ||||| ||||| ||| |||
392 TGACCAAAACCAAGGCCAGCGCCAAAACCAAGACCCCGCGCTAGCAGAA 441
326 ThrLysLysAlaGluGlyArgThrProAspAlaGluSerSergLysLys 342
      ||| ::|::: |||||:::
442 ACCCCATCACCCGAGGCTTATACACCAAGAGAACAATCAATGAGACAGA 491
```

342 sAlaProProGluAspSerGluAspAspMetGlnAlaGluAlaSerGlyc 359
492 TGAGTCTGCTGAACCAAGCTAGCGACGGTGTCTCAGGATGAGCATGTGGGTG 541
359 luAsnProAlaAlaLeuProGluAspAspGluValProGluAspThrGlu 375
542 AGCAA.....ATGACACCGCAT 558
376 HisAspAspProAsnSerAspProAspTyrTyrAsnAspMetProAlaVa 392
559 AAAGATGATCCT..... 570
392 lIleProValGluGluThrThrLysSerSerAsnAlaValSerMetProI 409
571TCCACAGGCATAGAGGTGTGCA 592
409 le.....PheAlaAlaPheValAlaCysAlaValAlaLeuVal 421
593 TCTGAGGAGAGATTATTATGTCATGATCGGTTCGCTATTGTGCTTTTG 639
seq_name: gb_est2:BE412830

seq_documentation_block:

LOCUS BE412830 954 bp mRNA linear EST 24-JUL-2000
DEFINITION MCG008.B08R990625 ITFC MCG Barley Leaf/Culm Library Hordeum vulgare
CDNA clone MCG008.B08, mRNA sequence.
ACCESSION BE412830
VERSION BE412830.1 GI:9410578
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 954)

REFERENCE
AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pechioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
International Triticeae EST Cooperative (ITFC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)

TITLE
JOURNAL
COMMENT Contact: Graner A
Institute for Plant Genetics & Crop Plant Research
Corrensstr. 3, D-06466 Gatersleben GERMANY
Tel: 49 39482 5521
Fax: 49 39482 5137
Email: a.graner@ipk-gatersleben.de
International Triticeae EST Cooperative (ITFC)
http://wheat.pw.usda.gov/genome.

FEATURES
source
1.954
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="MCG008.B08"
/clone_lib="ITFC MCG Barley Leaf/Culm Library"
/tissue_type="leaf/culm"
/dev_stage="etioloated"

BASE COUNT 266 a 192 c 259 g 227 t 10 others
ORIGIN

alignment_scores:
Quality: 111.00 length: 283
Ratio: 0.888 Gaps: 13
Percent Similarity: 44.170 Percent Identity: 22.615

alignment_block:
US-09-994-064-11 x BE412830 ..

Align seg 1/1 to: BE412830 from: 1 to: 954

151 TrpPheTyrValIleLysGlyAspAspGlyGluAspLysTyrCysProI 167
167 eTyrArgLysGluTyrArgGlyCys.GlyAspValGlnLeuSerGlu 183
63 CTATGTTGCAAACTAGAACTGAAAGCTTGCGGCTTATCGAAA 112
184 CysAlaValGlnSerAlaGlnMet.....TrrpAl 193
113 TCGGCTACAAAGAGTGGACAGAAAGAGGTCGGCTCTTGAGCAATTGGTG 162
193 aValAspTyrValProSerThrLeuValSerArgAsnGlyAlaGlyLeuT 210
163 T..... 163
210 hrIlePheSerProThrAlaAlaLeuSerGlyGlnTyrLeuLeuThrLeu 226
164ACTGCATCCGCAGTATTAGAGAGGCTGCATTTGCTGTGAC 205
227 LysIleGlyArgPheAlaGlnThrAlaLeuValThrLeuGluValAsnAs 243
206 AAA.....AGTTGATCACATA.GACATATCAGA 233
243 pArgCysLeuLysIleGlySerGlnLeuAsnPheLeuProSerLysCysT 260
234 GAAG..CAAAAGTTGTTAATGAGTGTCTG....GATGCAGAGAGCT 274
260 rPThrThrGlnGlnTyrGlnThrGlyPheGlnGlyGlnHisLeuTyrPro 276
275 GGCTAATGAGAGAAAGCAGCAG..... 297
277 IleAlaAspThrAsnThrArgHisAlaAspAspValTyrArgGlyTyrG 293
298 ...CAAGATGCTCTACCGAAGCATGCTAATCCTGTCTCTCT..GTTTC 341
293 uAspIleLeuGlnArgTrrpAsnAsnLeuLeuArgLysLysAsnProSera 310
342 TGACATTAAGAGAGAGGCTGAAGCACTTGACAGGTTCTGCAAAACCATCA 391
310 lAprAspProArgProAspSerValProGlnGluIlePro..AlaVal 325
392 TGACCAAAACCAAGCCAGCGCCAAACACAGACACCCCGGCTAGACGAA 441
326 ThrLysLysAlaGluGlyArgThrProAspAlaGluSerSerGlyLys 342
442 ACCCATCACCCGAGGCTTATACACACAGAAACAATCAATGAGAGAGA 491
342 sAlaProProGluAspSerGluAspAspMetGlnAlaGluAlaSerGlyc 359
492 TGAGTCTGCTGAACCAAGCTAGCGACGGTGTCTCAGGATGAGCATGTGGGTG 541
359 luAsnProAlaAlaLeuProGluAspAspGluValProGluAspThrGlu 375
542 AGCAA.....ATGACACCGCAT 558
376 HisAspAspProAsnSerAspProAspTyrTyrAsnAspMetProAlaVa 392
559 AAAGATGATCCT..... 570
392 lIleProValGluGluThrThrLysSerSerAsnAlaValSerMetProI 409
571TCCACAGGCATAGAGGTGTGCA 592
409 le.....PheAlaAlaPheValAlaCysAlaValAlaLeuVal 421
593 TCTGAGGAGAGATTATTATGTCATGATCGGTTCGCTATTGTGCTTTTG 639
seq_name: gb_gss:CNS06X9U
seq_documentation_block:

LOCUS CNS06X9U 955 bp DNA linear GSS 06-JUL-2001
DEFINITION T3 end of clone AX0AA039F09 of library AX0AA from strain CBS 7064
ACCESSION AL19464
VERSION AL19464.1 GI:12202642
KEYWORDS GSS.
SOURCE Pichia farinosa.
ORGANISM Pichia farinosa
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE
AUTHORS de Montigny,J., Spehner,C., Souclet,J., Tekala,F., Dujon,B.,
Wincker,P., Artiguenave,F. and Potier,S.
TITLE Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
sorbitophila
JOURNAL FEBS Lett. 487 (1), 87-90 (2000)
MEDLINE 20584725
REFERENCE
AUTHORS Souclet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Florente,B.,
Malpertuy,A., Neugeglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
REFERENCE
AUTHORS 3 (bases 1 to 955)
TITLE Genoscope.
JOURNAL Direct Submission
MEDLINE Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES
source
1. 955
location/Qualifiers
/organism="Pichia farinosa"
/strain="CBS 7064"
/db_xref="taxon:4920"
/clone="AX0AA039F09"
/clone_lib="AX0AA"
/note="end : T3"

BASE COUNT 177 a 219 c 235 g 312 t 12 others
ORIGIN

alignment_scores:
Quality: 111.00 Length: 186
Ratio: 1.220 Gaps: 8
Percent Similarity: 48.925 Percent Identity: 25.806

alignment_block:
US-09-994-064-11 x CNS06X9U/rev ..
Align seg 1/1 to reverse of: CNS06X9U from: 1 to: 955

296 LeuGlnArgTrpAsnAsnLeuLeuArgLys.LysAsnProSerAlaProA 312
||||:|||| :|||:|||| :|||:||||:||||:||||: |||:
799 CTTAGAAGAACCACTTAGAAGATCCGGAAGAACCAAGATCCAG 750
312 sPProArgProAsp.....SerValProGln..... 320
::: ||||| ::| |||:::
749 AGGAACCAACCACTTAGAAGAACCAAGAACCAAGATCGGAAGAGCYA 700

321GluIleProAlaValThrLysLysAlaGluGlyAr 332
||||:|||| :|||:|||| :|||:||||:||||:||||: |||:
699 CTGGAAGAAGAACCACTTACCAAGAGCTGCCAGAGCTGCCAGAGTCGGA 650
332 gThrProAspAlaGluSerSerGluLysLysAlaProProGlu.....A 347
: ||| |||::| ||| : ||||| :
649 AGAGCCGCTGGAAGAAGAACCAAGAGCTGCCAGAGCTTACCAAGCTTCAG 600
347 sPserGluAspAspMetGlnAlaGluAlaSerGlyGluAsnProAlaAla 363
:|||||::| :|||: ||| :|||:||||
599 AGTCAGAAGAACCACCTGGAAGATGAGCCACCACCTGGAAGATCCA..GAG 553
364 LeuProGlu.....As 367
||||||| ||
552 CTACCAAGACCAAGACCACTGGAAGATGAACCACCAACCGCTAGATGA 503
367 pAspGluValProGluAspThrGlu..... 375
| |||:|||||::| |||
502 TCCAGAGCTACCAAGAGTCGGAAGAACCGCTAGAAGAAGACCACTTCAC 453
376HisAspAspProAsnSerAspProAsp 384
:||||| :|||:||||| :|||:|||||: |||:
452 TAGAGGAGCCACCGCTAGAACCCGACAGAGGATCCACCGCGGAACAGAA 403
385 TyrTrpAsnAspMetProAlaValIleProValGluGluThr..... 398
::||| ||||| |||:|||||: |||:
402 GATCCAGAAGATCCACCGGACGACGACCAACCGCTAGAGATCAACCAACG 353
399ThrLysSerSerAsnAlaValSerMetProIlePheAlaAla 413
||||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
352 TCTCTTGACGAGACCTGGAGCGACGCTCTCATTTCTTATCAACTCAGCGT 303
413 heValAlaCysAlaValAlaLeuValGlyLeuLeuValTrpSerIleVal 429
:||||:||||:||||: ||| :|||: |||: |||: |||: |||: |||:
302 TGACAGCAGCAGCAACAGCTAAGGTAATTGCAACGTTAGAAAACCTCAT 253
430 LysCys 431
|||
252 TTGTGY 247

seq_name: gb_est2:BE599504

seq_documentation_block:
LOCUS BE599504 533 bp mRNA linear EST 18-AUG-2000
DEFINITION P11_88_C12.b1_A002 Pathogen induced 1 (P11) Sorghum bicolor cDNA,
mRNA sequence.
ACCESSION BE599504
VERSION BE599504.1 GI:9854577
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
AUTHORS 1 (bases 1 to 533)
Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 515
POLYA=NO.

Align seg 1/1 to: BE599342 from: 1 to: 546

```
284 HisAlaaspValTyrArgGlyTyrGluAspIleLeuGlnArgTrpAs 300
    ||| ::| ||| ::| ::| ::| ::| ::| ::| ::|
25 CATATGAATGAACTTTACATCTGAGATCAGAAATTTGATTAATAGG..... 69

300 nasnleuLeuArgLysLysAsnProSerAlaProAspProArgProAsps 317
    ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
70 .ATTCCTTACCAAAAGCAAAATGAACACTGATCCCTGAGAACAAAGATGATG 118

317 ervaProGlnGluIleProAlaValThrLysLysAlaGluGlyArgThr 333
    ::| ||| ::| ::| ::| ::| ::| ::| ::|
119 GA.....GAATCTGATGATGATTAATGATGATGATGATGAGACGAT 159

334 ProAspAlaGluSerSerGluLys..... 341
    ||| ||| ||| ||| ||| ||| ||| ||| |||
160 GAGGATGCTGAAAAACCAAGAAAGATGATGCTGTCGACGAGGATCAGA 209

342 .....LysAlaProProGluAspSerGluAspAspMetGlnA 354
    ::| ::| ::| ::| ::| ::| ::| ::| ::|
210 TGACAAATGGGAATGAGGAGGAGATGACGACGACGATGATGACGATC 259

354 laGluAlaSerGlyGluAsnProAlaAlaLeuProGluAspAspGluVal 370
    ||| ||| ::| ::| ::| ::| ::| ::| ::|
260 CTGAAGCTAATGCTGAAGAGGAAGC.....GATGATGATGATGATGCT 303

371 ProGluAspThrGluHisAspAspProAsnSerAspProAspTyrTrpAs 387
    ||| ||| ||| ::| ::| ||| ||| ::| ::| |||
304 GGAGAGGATGAAGAGGAGGATGACGACGATGACGACGGGACGGTGACAA 353

387 naspmetProAlaValIleProValGluGluThrThrLysSerSerAsnA 404
    ||| ||| ||| ||| ||| ||| ||| |||
354 TGACGAGAGATGAT.....GAGGAAGAAGAGAAGAGGATGACGATG 397

404 laval.....SerMetProIlePhe 410
    ||| ||| ||| ||| ||| ||| ||| |||
398 ACGTCCCAACCACTACTAAGAAGAGAAATGATCGATGCCCTGTTCCG 447

411 ..AlaAlaPheValAlaCysAlaValAlaLeuValGlyLeuLeuValTr 426
    ::| ||| ||| ||| ||| ||| ||| ||| |||
448 CCTGCGCAGCAGTATCTGTTTGTGGTGGTGG..GGGTTTTTGTGCTTG 494

426 pSer 427
    ||| |||
495 GTCG 498
```